Roles of the Histone Chaperone FACT in Drosophila Gene Regulation and Chromatin Architecture

Thesis

Roles of the Histone Chaperone FACT in Drosophila Gene Regulation and Chromatin Architecture

Theophilus Tettey, M.Sc.

A thesis submitted in fulfillment of the

requirements of the Open University

for the Degree of Doctor of Philosophy

The Stowers Institute for Medical Research

Kansas City, USA

an Affiliated Research Center of the

Open University, UK

June 2017

Dedication

To my late dad, David Tettey

my mum, Naomi Nettey,

Sisters, Eunice and Lilly

and,

daughter, Anneliese

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisors, Dr. Joan W. Conaway, Dr. Ronald C. Conaway and Dr. Marco Blanchette for their enormous support through my PhD study. Their patience, guidance, constructive criticisms, advice, immense knowledge and contributions have helped me greatly throughout my studies. They have provided a well-resourced laboratory with an intellectually stimulating environment. Without them my PhD dream would have forever been buried in my wildest dreams. I could not have imagined having better mentors for my research work.

Sincere thanks also to the rest of my thesis committee members, Dr. Jerry Workman and Dr. Peter Baumann, for taking time off to attend all my committee meetings and providing constructive criticisms and useful suggestions.

I would also like to thank the coordinator of the OU program and my third-party monitor, Dr. Leanne Wiedemann, for playing an instrumental role in my PhD study. She always made sure we met deadlines and did all necessary paperwork in a timely manner. Her office was always opened to discuss anything from science to topics on professional development goals. She is a fantastic coordinator. Also, special thanks to the former Open University administrative assistant, Shelly Hornbuckle, and the current administrative assistant, Lisa Hodges for their support in scheduling all my meetings.

I would also like to thank all past and current members of my laboratory for creating a friendly atmosphere that fosters collaborations, critical thinking and innovation. Lab meetings have always been good times to discuss science and receive constructive criticisms. Special thanks to Dr. Chieri T. Sato and Dr. Shigeo Sato who were always willing to share reagents, protocols and ideas to support my research work. I would also like to thank Dr. Kai Chen for showing me how to perform chromatin

immunoprecipitation followed by sequencing experiments which later became a central part of my research work.

Special thanks to Xin Gao, Benjamin Story, Alex Chitsazan, Dr. Chris Seidel and Dr. Marco Blanchette for analyzing the sequencing data. Without them the data would have been impossible to analyze.

Special thanks to Dr. John Lis and Dr. Hojoong Kwak for sharing their lab PRO-seq (Precision nuclear run-on and sequencing) protocol with me. Dr. Kwak was very helpful in providing free advice to get the experiments working. I would also like to thank Dr. Robert L. Glaser for sharing his anti-H2Av serum with me to perform ChIP experiments. Thanks also to Dr. Julia Zeitlinger and Wanqing Shao for sharing their Pol II (Rpb3) antibody.

Thanks also to my undergraduate mentor Dr. Yaa D. Osei for encouraging me to pursue a PhD.

Final thanks to my family and friends for their prayers and moral support throughout my study.

Abstract

The highly conserved histone chaperone FACT (Facilitates Chromatin Transcription) is thought to contribute to the disassembly and reassembly of nucleosomes in the wake of RNA polymerase II passage through chromatin. In yeast, mutation of FACT subunits leads to decreased nucleosome occupancy at the promoters and transcribed regions of genes. In addition, genetic experiments suggest that yeast FACT plays a critical role in restricting initiation of transcription to promoters, ensuring that only appropriate mRNAs are synthesized.

However, FACT's roles in chromatin biology and transcriptional regulation in higher eukaryotes are not well understood. Using Drosophila S2 cells as a model, I observe that depleting levels of the FACT complex results in aberrant transcriptional regulation of approximately 20% of expressed genes. We hypothesized that FACT-dependent alterations in transcription occur because nucleosome disassembly and reassembly are defective in FACT deficient S2 cells. To address that hypothesis, I monitored the nucleosome positioning using genome-wide micrococcal nuclease protection assays, and the distribution of bulk histones, histone variants, and histone modifications using ChIP-seq. I observed that FACT depletion has little effect on total nucleosome occupancy or positioning or on the distribution of either the histone variant H2Av or the histone 3 lysine 56 acetylation (H3K56ac) mark. On the other hand, the loss of FACT alters the distribution of other histone modifications, including histone 3 lysine 4 trimethylation (H3K4me3) and histone 3 lysine 36 trimethylation (H3K36me3), in the promoters and transcribed regions of genes. In particular, H3K4me3, a mark of active promoters, was reduced at promoters and increased in gene bodies of the FACT depleted cells. Furthermore, levels of H3K36me3, a mark of active transcription over transcribed regions, was reduced in gene bodies but increased beyond the polyadenylation site. The levels of total (Rpb3) and transcriptionally engaged Pol II (PRO-seq) were reduced at promoters upon the loss of FACT.

Taken together, our results are consistent with the model that FACT contributes to the interplay between chromatin architecture and control of promoter-proximal pausing.

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

ATP	Adenosine triphosphate
bp	Base pair
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CTD	C-terminal domain
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dsRNA	Double-stranded RNA
GO	Gene ontology
ml	Milliliter
mM	Milimolar
mRNA	messenger RNA
ncRNA	Noncoding RNA
ng	Nanograms
nt	Nucleotide
oligos	Oligonucleotides
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline with Tween-20
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
Pol II	RNA polymerase II
RNA	Ribonucleic acid
RNAi	RNA interference
RNA-seq	RNA sequencing
RPKM	Reads per kilobase of million mapped reads
RT-PCR	Reverse-transcription polymerase chain reaction
SDS	Sodium dodecyl sulfate
snoRNA	small nucleolar RNA
snRNP	Small nuclear ribonucleoprotein particle
T7	Bacteriophage T7 promoter
tRNA	Transfer RNA
UTR	Untranslated region
UV	Ultraviolet
µg	Microgram
µl	Microliter
µM	Micromolar

Chapter 1

Introduction

1.1 Chromatin and gene regulation

The term “chromatin” was initially coined by W. Flemming to describe the substance in the cell nucleus that is refractile in nature and can be readily stained (Flemming, 1882). By the beginning of 1970s, more progress had been made to understand the chromatin structure. At that time, evidence suggested that histones reduce RNA transcription significantly (Allfrey et al., 1964), but not much was known about the nature or relative amounts of individual histone proteins. In 1967 E. W. Johns used polyacrylamide gel electrophoresis to quantify the five major histone fractions, f1, f2(a)1, f2(a)2, f2(b) and f3 prepared from calf thymus (Johns, 1967). These "fractions" are now known as histones H1/H5, H2A, H2B, H3, and H4 respectively.

Work by Hewish and Burgoyne provided evidence for the existence of a repeating chromatin sub-structure with regularly distributed sites that can be cleaved by magnesium- and calcium-dependent endonuclease. They observed a non-random ladder of DNA after digestion of rat liver chromatin by a nuclear calcium and magnesium dependent endonuclease. Based on this result they proposed that, “chromatin has some simple, basic, repeating substructure with a repetitive spacing of sites that are potentially accessible to the Ca-Mg endonuclease” (Hewish and Burgoyne, 1973).

To study the structure of chromatin, Olins and Olins visualized chromatin isolated from fresh rat thymus, rat liver, and chicken erythrocytes using electron microscopy. The micrographs they obtained show that chromatin fibers appear as “particles on a string” (Olins and Olins, 1974). A similar observation was made by Woodcock and colleagues in micrographs of chromatin isolated from a wide range of tissues including chicken

erythrocytes, amphibian erythrocytes, chlamydomonas and fruit fly salivary glands. These observations led them to hypothesize that the “spherical particles are a fundamental structural unit of chromatin”(Woodcock et al., 1976).

Based on results from biochemical and x-ray diffraction experiments, Kornberg and Thomas proposed that the chromatin structure is composed of repeating units of eight histones consisting of a histone H3-histone H4 tetramer, dimers or oligomers of histones H2A andH2B, and about 200 bp DNA (Kornberg, 1974; Kornberg and Thomas, 1974). Evidence for the proposed tetramer came from cross-linking and sedimentation experiments followed by electron microscopy.

Oudet and colleagues were the first to introduce the term “nucleosome” to describe the particles that aligned along the chromatin fiber. They chose the term nucleosome because of the nuclear origin of the particle and its similarity to the “nu” (ν)-bodies described by Olins and Olins. Based on biochemical and electron microscopic examination, Oudet and colleagues showed that “ the fundamental structure of chromatin fibers is composed of a flexible chain of spherical particles with a diameter of 124-130 Å connected by DNA filaments,” consistent with Kornberg’s model of the nucleosome structure (Oudet et al., 1975).

In addition, Finch and colleagues prepared nuclei from rat liver and subjected the chromatin to micrococcal nuclease digest. They then visualized the nucleosome core particle crystals using X-ray diffraction and electron microscopy. They observed that the core particle is flat with dimensions of about 110 x 110 x 57 Å, and they proposed it is divided into two layers with 1 ¾ super helical turns of DNA on the histone octamer (Finch et al., 1977). In order to have a better resolution of the electron density map of the core particle from 20 Å, Richmond and colleagues improved the preparation of the histone core crystals. They then solved the nucleosome core particle to a 7 Å resolution

(Richmond et al., 1984) which confirmed the dyad . The resolution of the nucleosome core particle was further improved by Luger and colleagues to 2.8 Å (Luger et al., 1997) and later refined to 1.9 Å (Richmond and Davey, 2003).

The current and prevailing model is that, in eukaryotic cells, DNA is tightly packaged into nucleosomes to form chromatin. Nucleosomes therefore form the basic repeating unit of the chromatin consisting of a histone octamer with 146 bp DNA wound 1.65 times around H3-H4 tetramer and two H2A-H2B dimers. Linker DNA of variable lengths connects the nucleosomes.

1.1.1 Histone modifications

In the 1960s, Allfrey and colleagues presented a groundbreaking observation suggesting roles of histones and histone modifications (acetylation and methylation) in transcription regulation. After being inspired by work from Stedman and Stedman, Allfrey and colleagues initially demonstrated that "histones inhibit RNA synthesis in isolated thymus nuclei" and that "removal of the histones from the nucleus results in an increased rate of messenger RNA synthesis" (Allfrey et al., 1964). Allfrey and colleagues further showed that methylating and acetylating histones relieves inhibition of RNA polymerase activity (Allfrey et al., 1964).

Today, there are vast array of enzymes identified capable of modifying certain residues on the globular core domains and histone tails. For example, the lysine residues on histone tails can be acetylated, SUMOylated, ubiquitinated, biotinylated and methylated. Arginine residues on histone tails can be methylated, citrullinated and ribosylated. Likewise, glutamic residues can undergo ADP-ribosylation. Serine and threonine residues can be phosphorylated. Proline residues can be isomerized. More recent evidence suggest that most of these modifications have biological implications

(Latham and Dent, 2007). For the purpose of this thesis, I will focus on histone modifications related to transcription.

1.1.1.1 H3K36me3

The lysine residue 36 on H3 tails can be mono- , di- , and tri-methylated, by a wide range of methyltransferases. Current literature describes at least 8 enzymes, all bearing a SET domain capable of methylating H3K36 in mammals (Wagner and Carpenter, 2012). The methylated lysine residues can also be enzymatically de-methylated depending on the need of the cell (Black et al., 2012).

Histone methyl transferases methylate the lysine residues on H3 by utilizing S-adenosylmethionine. Set2 was initially identified in yeast as a histone methyl transferase that mediates tri-methylation of lysine 36 of the histone H3 tail (Strahl et al., 2002). Genome-wide data obtained from ChIP of H3K36me3 coupled with DNA microarray suggests (i) that H3K36me3 occupies gene bodies, with a bias towards 3'ends of open reading frames and (ii) that it correlates with transcriptional activity (Pokholok et al., 2005).

Studies from the Workman lab and others have shown that the presence of H3K36me3 across gene bodies recruits Rpd3S deacetylase, which removes acetyl residues across gene bodies rendering them in a hypo-acetylated state thereby preventing cryptic transcription (Carrozza et al., 2005). Biochemical and genetic experiments reveal that the H3K36me3 methyltransferase Set2 interacts with Ser2 phosphorylated Pol II through the Set2 WW domain and plays a role in transcription elongation (Li et al., 2002). The model is therefore that Set2 is recruited to Ser2 phosphorylated Pol II and that it methylates histone tails during transcription elongation.

The highly conserved JmJC domains of the lysine demethylases KDM4A, KDM4B, KDM4C and KDM4D mediate the demethylation of H3K36me3 by utilizing α-ketoglutarate in the presence of the cofactors Fe (II) and oxygen (Black et al., 2012; Shi and Whetstine, 2007).

Results of co-immunoprecipitation experiments performed in SETD2-depleted cells suggests H3K36me3 recruits the histone chaperone and elongation factor FACT to transcribed regions (Carvalho et al., 2013). Based on this observation, it has been proposed that Set2 regulates recruitment of FACT through H3K36me3.

1.1.1.2 H3K4me3

The highly conserved histone H3 tail residue lysine 4 can also undergo post-translational modifications. H3K4 can be mono-, di-, and tri-methylated by a collection of histone methyl transferases. Depending on the needs of the cell, methylation on lysine 4 residue of histone H3 tails can be reversed by histone demethylases.

Genome-wide mapping of H3K4me3 using chromatin immunoprecipitation followed by DNA microarray suggests that histone H3 lysine 4 tri-methylation is enriched at promoter regions of active genes (Pokholok et al., 2005). In addition, Pokholok and colleagues observed that levels of H3K4me3 occupancy correlate with transcriptional activity. Quite recent work using genomics suggests that the broadening of H3K4me3 into gene bodies is strongly associated with transcription elongation (Chen et al., 2015b).

Tri-methylation of H3K4 is carried out by the Set1/COMPASS, a highly conserved protein complex from yeast to humans (Briggs et al., 2001; Noma and Grewal, 2002; Shilatifard, 2012). H3K4me3 methylase Set1 has been shown to be associated with the 5' end of genes and interacts with Pol II phosphorylated on Ser5 of the Rpb1 CTD but not Pol

II phosphorylated on CTD Ser2 residues (Ng et al., 2003). KDM5A, KDM5B, KDM5C and KDM5D demethylates tri-methylated H3K4 (Black et al., 2012).

Stanlie and colleagues demonstrated connections between FACT and H3K4me3 in in experiments where levels of FACT were manipulated followed by ChIP-qPCR of H3K4me3. They observed that the reduction of FACT levels leads to a reduction of H3K4me3 levels over immunoglobulin μ and α heavy chain switch regions (Sμ and Sα, respectively), suggesting that FACT-dependent changes in H3K4me3 plays a role in Ig class switch recombination (Stanlie et al., 2010).

1.1.1.3 H3K56ac

H3 lysine 56 acetylation is one of the few modifications that occurs on the globular core domains of histones. Efforts from a number of researchers have demonstrated that H3K56 acetylation is well conserved from yeast to humans (Das et al., 2009; Ozdemir et al., 2005; Yuan et al., 2009). In yeast, H3K56 acetylation mostly marks newly synthesized histone H3 (Masumoto et al., 2005). Also, H3K56 acetylation increases during S-phase of the cell cycle. (Masumoto et al., 2005; Recht et al., 2006). Acetylation of lysine 56 also functions to maintain genome integrity (Yuan et al., 2009). In order to maintain an optimal pool of H3K56 acetylation in the cell, histone methyl transferases and deacetylases play an instrumental role regulating H3K56ac homeostasis. The yeast H3K56 acetyl transferase, Rtt109, was discovered after a collection of 4700 yeast mutants were screened for H3K56 acetylation using specific antibodies (Han et al., 2007). Orthologues of the H3K56 acetyl transferase, CBP and CBP/p300 have been found in Drosophila and humans respectively. On the other hand, yeast Hst3 and Hst4, *Drosophila* Sir2, human SIRT1 and SIRT2 mediate the deacetylation of H3K56ac (Celic et al., 2006; Das et al., 2009; Maas et al., 2006). Genome-wide data obtained from the ChIP of H3K56ac followed by

microarray analysis suggests, H3K56ac occupies predominantly promoters of active genes (Rufiange et al., 2007).

1.1.1.4 H2A.Z

H2A.Z is a histone 2A variant that is evolutionarily conserved from yeast to humans (Iouzalen et al., 1996). The H2A.Z amino acid sequence is about 60% identical to the canonical H2A histone (Jackson et al., 1996). Similar to histone H2A, H2A.Z can undergo a number of post-translational modifications including acetylation, methylation, ubiquitination and SUMOylation (Sevilla and Binda, 2014).

Histone H2A.Z has been shown by a number of labs to be essential for survival. Mutations in H2A.Z are lethal in a wide range of organisms including the fly, frog, mouse, and tetrahymena but are not lethal in yeast (Bonisch and Hake, 2012). However mutations in yeast H2A.Z result in slow growth (Jackson and Gorovsky, 2000).

Genome-wide experiments have shown that H2A.Z is preferentially enriched at promoters of genes from yeast to humans (Guillemette et al., 2005). H2A.Z is associated with both active and inactive genes depending on the organism. In yeast, H2A.Z occupancy at promoters is poorly correlated with transcriptional activity (Raisner et al., 2005). In contrast, Barski and colleagues showed using data from ChIP-seq of H2A.Z in human $CD4^+$ T cells that H2A.Z occupancy correlates with level of gene expression (Barski et al., 2007). The yeast SWR1 complex, a member of the Swi2/Snf2 family of ATPases, was the first enzyme demonstrated to mediate the exchange of histone H2A-H2B dimers for H2A.Z-H2B dimers in an ATP-dependent manner (Krogan et al., 2003; Mizuguchi et al., 2004). Tip60, the Drosophila homologue of the Swrc-1 complex, mediates the deposition of H2Av (Kusch et al., 2004). INO80 on the other hand mediates the eviction of H2A.Z in an ATP dependent manner (Papamichos-Chronakis et al., 2011).

Recently, Jeronimo and colleagues have demonstrated that FACT mutations are associated with mislocalisation of H2A.Z into gene bodies in yeast and an increase in cryptic transcription (Jeronimo et al., 2015). Based on this observation, they proposed that the presence of H2A.Z within genic locations results in a less stable nucleosome leading to cryptic transcription.

1.2 Histone chaperones

Transcription, the making of RNA by RNA polymerase using a template DNA, occurs in a chromatin context *in vivo*. Transcription *in vitro* has been shown to be limited by the presence of nucleosomes. Nucleosomes therefore act as barriers that deny access to transcription factors and other transcription machinery and, as a consequence, play a major role in regulating gene expression. Cells have therefore evolved a collection of enzymes, examples of which are histone chaperones, that help alleviate the inhibitory role of nucleosome to allow transcription initiation and elongation.

In 1978, R. Laskey and colleagues coined the term 'molecular chaperone' to describe the activity of a chromatin assembly factor in Xenopus egg homogenate that prevents mixtures of DNA and histone from precipitating under physiological ionic strength (Laskey et al., 1978; Laskey et al., 1977). Further characterization of the chaperone reveal that it prevents inappropriate ionic interactions between DNA and histones. Histone chaperones, as they are currently termed, are defined as a family of proteins that bind histones and mediate one or more of the following processes; transport, storage, assembling, and disassembling of histones. Histone chaperones can have redundant and specific roles. Histone chaperones can bind preferentially to H2A-H2B histones, H3-H4 histones, or histone variants (Burgess and Zhang, 2013).

Histone chaperones are involved in distinct steps in the assembling of nucleosomes. First, after histone proteins are synthesized in the cytosol, they are imported into the nucleus. Some histone chaperones, such as Nap1, help to shuttle newly synthesized histones from the cytoplasm to the nucleus in part through regulation of the importin-histone interaction. Second, a soluble pool of histones must be maintained to meet challenges during stress conditions and histone chaperones such as NASP act as reservoirs to maintain histone supply. Third, some histone chaperones and histone-binding proteins, such as RbAp46 and Asf1, directly regulate the enzymatic activity of histone-modifying enzymes. Fourth, other histone chaperones are thought to be directly involved in the eviction and deposition of histones during transcription or replication.

1.2.1 FACT complex

1.2.1.1 Discovery of FACT

The FACT complex is evolutionarily conserved from yeast to humans. FACT was discovered as CP complex (Cdc68 and Pob3) and SPN (Spt16, Pob3, and Nhp6) in yeast, DUF (DNA unwinding factor) in Xenopus, and FACT (facilitates chromatin transactions, composed of SPT16 and SSRP1) in humans. It is therefore not surprising that FACT is implicated in transcription, replication and DNA repair pathway.

The yeast Spt16 (Suppressor of Ty insertion) was identified through yeast genetics as a high-copy-number suppressor of δ insertion mutations in the 5’ regions of the *HIS4* and *LYS2* genes of *S. cerevisiae* (Clark-Adams et al., 1988). Through genetic mapping and complementation analysis it was demonstrated that SPT16 is identical to CDC68 (cell division cycle) (Malone et al., 1991). Further characterization showed yeast Spt16 is important for growth and plays a role in transcription regulation (Malone et al., 1991). Brewster and colleagues discovered through purification that Cdc68 exists as a

heterodimer with Pob3, which they termed the CP complex (Brewster et al., 1998). Prior to the discovery of the CP complex, work from Wittmeyer and colleagues had demonstrated that the yeast DNA polymerase 1 (DNA Pol 1) protein interacts with Cdc68/Spt16 and Pob3 (Wittmeyer and Formosa, 1997), implicating yeast FACT in DNA replication.

SSRP1 (in Xenopus known as DUF87) is the metazoan orthologue of yeast Pob3. SSRP1 has a highly conserved HMG1 motif that confers DNA binding ability. For this reason, Formosa and colleagues (2001) considered the possibility that an HMG1 containing protein might interact with the yeast CP complex. Through genetic screening, the HMG1 containing protein Nhp6 was found to display a strong interaction with the CP complex. Purified Nhp6 is capable of binding to nucleosomes, and Spt16-Pob3 binds to Nhp6-bound nucleosomes. Based on this observation, Formosa and colleagues proposed that Nhp6 provides a "landing pad" for Spt16-Pob3 on nucleosomes. Binding to nucleosomes of Spt16-Pob3 and Nhp6 (collectively referred to as the SPN complex) can result in nucleosome remodeling (Formosa et al., 2001).

Human FACT was identified in Reinberg's lab as an activity that stimulates transcription elongation over chromatinized template *in vitro* (Orphanides et al., 1998). The FACT activity copurified with polypeptides that migrate at 140 (p140) and 80 kDa (p80) on a PAGE gel. Determination of p80's sequence revealed that it is identical to a protein that had been identified as SSRP1 (Orphanides et al., 1999). SSRP1 was initially discovered during a search for cellular factors that can process DNA modified by the highly potent antitumor drug, cis-diamminedichloroplatinum or cisplatin (Toney et al., 1989), which has been shown to disrupt DNA replication and transcription by altering the structure of the DNA. Further characterization of the structure of SSRP1 led to the

discovery that it includes peptide sequences similar to HMG box 1 and 2 proteins (Bruhn et al., 1992).

Drosophila FACT, also known as dFACT, is a heterodimer of dSpt16/Dre4 and dSSRP1. dFACT was identified through a purification to identify GAGA factor associated proteins. Anti-flag purification performed with nuclear extracts prepared from embryos of transgenic flies expressing Flag-tagged GAGA factor identified two proteins on the stained SDS-PAGE gel in addition to GAGA factor. Peptide sequencing followed by western blotting showed that these proteins were dSpt16 and dSSRP1. Drosophila FACT has been shown to modulate chromatin structure *in vitro* and to function with GAGA factor in regulation of *Hox* gene expression (Shimojima et al., 2003).

Okuhara and colleagues used old fashioned chromatography to purify DUF (DNA unwinding factor) from Xenopus egg extract as an activity that introduces negative supercoils into a relaxed duplex DNA in the presence of topoisomerase I. Okuhara and colleagues monitored the activities of eluate fractions from the six successive column purification of Xenopus egg extract capable of introducing negative supercoils into duplex DNA in the presence of topoisomerase I. The DNA unwinding activity migrated at 140 kDa and 87 kDa and eluted as a single band (250 kDa) through gel filtration chromatography. Peptide sequencing followed by sequence analysis of the DUF-encoding cDNAs showed that DUF140 and DUF87 are homologous to yeast Cdc68/Spt16 and human SSRP1 respectively. In this same study, DUF was found to play a major role in DNA replication. Immunodepletion of DUF from Xenopus extracts using anti-DUF140 antibodies reduced their ability to replicate added sperm chromatin to 5- 25%. In contrast, DNA replication was restored to about 75% of the control when purified DUF was added back to the reaction (Okuhara et al., 1999).

1.2.1.2 Structural insights

In metazoans, the FACT complex is a heterodimeric protein complex consisting of Spt16 (Suppressor of Ty 16) subunit and SSRP1 (Structure specific recognition protein 1) with molecular weights of ~140 kDa and ~80 kDa respectively (Orphanides et al., 1999). However, the yeast analogue of the SSRP1 subunit is a complex consisting of Pob3 (Pol I-binding protein 3) and Nhp6 (Non-histone protein 6) (Brewster et al., 2001; Formosa et al., 2001).

Work from a number of labs has defined structural domains of the FACT complex. The Spt16 subunit is composed of the N-terminal domain (NTD), the dimerization domain (DD), the middle domain (MD) and the C-terminal domain (CTD) (Keller and Lu, 2002; O'Donnell et al., 2004; VanDemark et al., 2006).

O'Donnell and colleagues demonstrated that the Spt16 NTD is well conserved and not essential for yeast viability. Purification of the Spt16NTD followed by crystallization showed that the Spt16 NTD has two lobes with the C-terminus lobe folded like "pita bread" and structurally looks like amino-peptidases (VanDemark et al., 2008). Using gel filtration and GST-pulldown assays Ladurner's lab showed that Spt16N interacts with the globular domains and tails of H3-H4 histones (Stuwe et al., 2008).

The Spt16 subunit of FACT dimerizes with SSRP1/Pob3 through the Spt16DD and SSRP1DD dimerization domains (Keller and Lu, 2002). Hondele and colleagues purified, crystalized and solved the structure of the dimerization domains of yeast FACT (yFACT: Spt16D-Pob3N). They observed that Sp16D contains a single pleckstrin homology-like (PHL) module and Pob3 contains tandem PHL modules with the second module lacking the capping helix. The yFACT dimerization domains does not interact with H2A-H2B histones nor H3-H4 but co-precipitates with the large subunit of DNA polymerase α.

The Spt16 middle domain (Spt16M) is evolutionarily conserved and it is composed of a tandem pleckstrin homology-like (PHL) modules (Kemble et al., 2013). Kemble and colleagues observed, using EMSA, that Spt16M contacts H3-H4 histones with low affinity and that the Spt16M domain does not contribute significantly to the binding of Spt16-Pob3 to H2A-H2B. Quite recent work by Hondele and colleagues using crystals of the Spt16 middle domain (Spt16M) purified from *Chaetomium thermophilum* reported that Sp16M interacts with a higher affinity with H2A-H2B than with H3-H4 (Hondele et al., 2013). An alternative report from Kemble and colleagues determined that the acidic Spt16CTD tails mediate the interaction between FACT and H2A-H2B rather than the middle domain (Kemble et al., 2015).

The Spt16 CTD is highly acidic and it interacts with nucleosomes and stimulates transcription elongation. Mutations of FACT bearing a Spt16 CTD deletion cannot assemble nucleosomes and does not stimulate transcription elongation (Belotserkovskaya et al., 2003).

The metazoan SSRP1 subunit of FACT is composed of an N-terminal domain (NTD), dimerization domain (DD), and middle domain (MD), as well as two intrinsically disordered regions flanking an high-mobility group HMG-1 domain (Tsunaka et al., 2009; VanDemark et al., 2006).

The roles of SSRP1 domains in nucleosome assembly and disassembly are not fully understood. Reinberg's lab demonstrated using biochemical assays that full length recombinant hSSRP1 binds to histone H3-H4 tetramers. In addition, SSRP1 HMG-1 domain mediates FACT binding to DNA in a non-sequence specific manner to enable effective disassembly and reassembly of chromatin (Masse et al., 2002). Apart from the SSRP1 HMG-1 domain, SSRP1-M is capable of binding to dsDNA. Zhang and colleagues obtained crystal structure of the SSRP1 middle domain at a 1.93 Å resolution. The structure

contains tandem pleckstrin homology domains and is positively charged on each side. They discovered through pulldown and gel filtration assays that SSRP1-M binds non-specifically to DNA but binds weakly, if at all, to histones (Zhang et al., 2015).

1.2.1.3 How FACT reorganizes the nucleosome

There are currently a number of models describing how FACT displaces/reorganizes nucleosomes to allow the transcription machinery to gain access to the underlying sequences in cells.

Reinberg's lab proposed the dimer displacement model based on observations that suggest that hFACT removes either one or both of the H2A-H2B dimers to allow RNA Pol II passage (Belotserkovskaya et al., 2003; Orphanides et al., 1999). Reinberg's lab observed that the hSpt16 and SSRP1 subunits bind H2A-H2B dimers and H3-H4 tetramers respectively. Also, incubating purified hFACT with an immobilized chromatinized template leads to loss of about half of H2A-H2B dimers from nucleosomes. In addition, FACT mediated transcription through chromatinized templates leads to the formation of a hexasome.

The nucleosome reorganization/ nucleosome breathing model (Hondele et al., 2013; Kemble et al., 2015; Xin et al., 2009) was initially proposed by Xin and colleagues who provided evidence that yFACT reorganizes nucleosomes into a looser and more dynamic form with the original histone composition. Structural work from Hondele and colleagues also supported this model with results that showed that the middle domain of Spt16 makes several contacts with histones and gradually disrupts H2A-H2B-DNA interaction enhancing "nucleosome breathing" (Hondele et al., 2013). More recent structural and biochemical work by Kemble and colleagues suggest that "FACT maintains

nucleosomes in an altered, loosened structure at least in part by competing with DNA for binding both H2A-H2B dimers in a nucleosome" (Kemble et al., 2015).

1.2.1.4 Role of FACT in transcription regulation

A variety of evidence supports FACT's role in regulating RNA transcription. As noted earlier, human FACT was initially identified and purified using biochemical methods as an activity that stimulates transcription elongation over chromatinized template *in vitro* (Orphanides et al., 1998). Mutations in genes that encode transcription elongation factors in yeast often cause growth defects in the presence of 6-azauracil (6-AU) (Powell and Reines, 1996). FACT was implicated in transcription elongation after yeast Spt16 mutants showed growth defects in the presence of 6-AU (Formosa et al., 2001; John et al., 2000; Orphanides et al., 1999). In addition, the observation that Spt16 interacts genetically and physically with elongation factors such as Paf1 complex members, TFIIS and Spt4 subunit of the Spt4/5 complex, further implicates FACTs in transcription elongation (Orphanides et al., 1999; Squazzo et al., 2002). Immunofluorescence staining of FACT and elongating RNA polymerase II on Drosophila polytene chromosome shows they colocalize at active sites (Saunders et al., 2003).

Crosslinking followed by ChIP-qPCR of FACT on active Pol II genes showed yeast FACT occupies transcribed regions (Mason and Struhl, 2003). In this experiment, Mason and Struhl showed that FACT is enriched on a number of active Pol II transcribed genes. Work from Ransom and colleagues demonstrated using yeast genetics and biochemistry that yFACT is involved in the disassembly of histone H2A/H2B from the PHO5 promoter to activate transcription, implicating FACT in transcription initiation (Ransom et al., 2009).

In addition, conditional inactivation of Spt16 in yeast results in decreased binding of TBP and TFIIB, components of the Pol II pre-initiation complex, at normal promoters

(Mason and Struhl, 2003), suggesting that FACT is necessary for TBP and TFIIB occupancy at promoters. On the other hand, mutations in yeast Spt16 result in transcription from cryptic promoters within the coding regions of some genes. This suggests that FACT plays a role in preventing inappropriate transcription initiations from promoter-like regions in cells (Kaplan et al., 2003; Mason and Struhl, 2003). In addition, *in vitro* studies show that Spt16 can mediate TBP binding to promoters (Biswas et al., 2005). These observations point to the role of FACT not only in transcription elongation but also in transcription initiation.

In a minimal transcription system, FACT can inhibit basal transcription of a supercoiled DNA template. On the other hand FACT, in concert with PTEFb, can promote transcription over naked DNA which was otherwise inhibited by NELF and DSIF (Wada et al., 2000).

In the quest to understand whether the role of FACT in transcription regulation is global or only specific to some genes, Li and colleagues purified RNA from human non-small cell lung carcinoma (H1299) cells depleted of either SSRP1 or Spt16 followed by spotted microarray analysis. The microarray data revealed that only ~1.3% of the 8308 genes tested were either up-, or down-regulated following the knockdown of SSRP1 and Spt16. In addition, Spt16 and SSRP1 regulate the expression of a common and unique sets of genes (Li et al., 2007). Thus the FACT complex does not only promote transcription elongation as it was originally discovered but also represses the expression of genes.

1.2.1.5 FACT's role in cancer

Over the years, the expression of FACT has been linked to a number of cancers including ovarian cancer, colon carcinoma, melanoma, pancreatic ductal adenocarcinoma, glioblastoma, hepatocellular carcinoma and neuroblastoma. Initial speculations of the

possible involvement of FACT in cancer came from work by Xiang and colleagues, where they monitored the expression of SSRP1 using northern blot and *in situ* hybridization in rat fetal kidney, adult kidney, and renal cell carcinoma induced by Fe-nitrilotriacetate. Northern blot analysis showed that SSRP1 was expressed in rat fetal kidney and renal cell carcinoma but not in the adult kidney. Further investigations using *in situ* hybridization led to the discovery that SSRP1 localizes in epithelial cells of fetal kidney, and neoplastic cells of renal cell carcinoma but not in cells of normal adult kidney, suggesting that SSRP1 might have a role to play in kidney development and carcinogenesis (Xiang et al., 1996). Subsequent work by Hudson and colleagues, using protein microarrays and autoantibodies from cancer patients, also identified SSRP1 as one of the proteins that is aberrantly expressed in ovarian cancer (Hudson et al., 2007). In addition, analyses of publicly available mRNA expression datasets showed that the expression of FACT is strongly associated with embryonic development, stem cells and differentiation, and oncogenes activity in mammals. Also, when C2C12 myoblasts were induced to differentiate into myotubes, the levels of FACT subunits were reduced, suggesting that the expression of FACT may be necessary to keep cells in an undifferentiated state (Garcia et al., 2011). Also, Koman and colleagues in their work with MMTV-neu transgenic mice, a model for breast cancer development discovered that SSRP1 was highly induced (Koman et al., 2012). FACT has been shown to drive metastasis in breast and lung cancer (Garcia et al., 2013). A recent report by Ding and colleagues implicated SSRP1 as a driver of metastasis in patients with hepatocellular carcinoma (HCC) (Ding et al., 2016). Ding and colleagues analyzed publicly available RNA expression data from patients with HCC and showed that levels of SSRP1 were increased. SSRP1 protein levels were also high in most HCC cell lines. In the bid to find out whether the high SSRP1 expression influence certain clinicopathologic factors in HCC patients, Ding and colleagues demonstrated that high

SSRP1 correlated with faster carcinoma enlarging and spreading. In addition, HCC patients with very high SSRP1 in their tumors have a shorter overall survival. In order to gain more insight into the role of FACT in HCC, levels of SSRP1 were manipulated in HCC cell lines. The knockdown of SSRP1 in HCC cells led to a significant increase in cell apoptosis. FACT therefore has become a therapeutic target to a class of anticancer drugs called curaxins (Di Bussolo and Minutolo, 2011; Gasparian et al., 2011) due to its implication in cancer.

Chapter 2

Role of FACT in gene expression

2.1 Introduction

A number of evidence supports FACTs role in regulating RNA transcription. As noted earlier, the human FACT was initially identified and purified using biochemical methods as an activity that stimulates transcription elongation over a chromatinized template *in vitro* (Orphanides et al., 1998). Crosslinking followed by ChIP-qPCR of FACT on active Pol II genes demonstrates that FACT occupies transcribed regions of specific genes (Mason and Struhl, 2003). yFACT is involved in the disassembly of histone H2A/H2B from the PHO5 promoter to activate transcription, implicating FACT in transcription initiation (Ransom et al., 2009).

Conditional inactivation of Spt16 in yeast results in decreased binding of members of the pre-initiation complex, TBP and TFIIB at normal promoters (Mason and Struhl, 2003). On the other hand, the mutations in Spt16 results in transcription from cryptic promoters within the coding regions of some genes. This suggests that FACT plays a role in preventing inappropriate transcription initiations from promoter-like regions in cells (Kaplan et al., 2003; Mason and Struhl, 2003). Relatively little is known about contributions of FACT to gene regulation in higher eukaryotes and, beyond evidence that FACT depletion leads to changes in expression of only a few genes in human cell line (Li et al., 2007), there is little information about the effects of FACT on gene expression genome-wide. Results of microarray analyses suggest that FACT depletion leads to either up- or down-regulation of only ~1.3% of the 8308 genes tested in human cells.

In this chapter, I begin to explore roles of FACT in global gene regulation in Metazoa, using Drosophila S2 cells as a model system. In initial experiments, I use

Chromatin immunoprecipitation to determine whether FACT is enriched at transcribed regions of the genome. In addition, I use RNA-seq to explore and define the consequences of SSRP1 and Spt16 depletion on transcript accumulation, genome-wide.

2.2 Materials and Methods

2.2.1 Cell culture

Drosophila S2 cells (GIBCO) were conditioned and maintained in Schneider's medium (GIBCO, # 21720) supplemented with 10% heat-inactivated fetal bovine serum (PAA, #A15-701) at 25°C prior to treatment with the dsRNA. All tissue culture manipulations were carried out in the tissue culture hood

2.2.2 Making double stranded RNA

2.2.2.1 Design of double stranded RNA

Double-stranded RNA (dsRNA) for RNAi was generated essentially as described (Blanchette et al., 2009); (Blanchette et al., 2005; Clemens et al., 2000). Briefly, the double stranded RNA (dsRNA) was generated from a PCR template carrying a T7 RNA polymerase promoter sequence (5'-TAATACGACTCACTATAGG-3') on each end. The primers for making the template begin with the T7 RNA polymerase promoter sequence 5'-TAATACGACTCACTATAGG-3' followed by a gene specific sequence of approximately 20-30 bp (see next section OR Table 1). For optimally effective knockdown, the dsRNA should be at least 400 bp long. In order to reduce the occurrence of false positive phenotypes, two non-overlapping dsRNA targeting the gene of interest were used for most RNAi experiments. As a negative control, dsRNA targeting eGFP (enhanced green fluorescent

protein) was generated. eGFP was used as control since it has either no sequence homology to the fly genome.

Table 2-1 Primer list for making T7 template

Primer name	5' - 3' sequence
T7-eGFP.fwd	TAATACGACTCACTATAGGGACCAGGAGCGCACCATCTTCT
T7-eGFP.rev	TAATACGACTCACTATAGGGAGGCGGTCACGAACTCCAGCA
T7-SSRP1a.fwd	TAATACGACTCACTATAGGGAGGAAGGGCGGCGAGATGTGGAAG
T7-SSRP1a.rev	TAATACGACTCACTATAGGGAGGAGGCCTCATCTTCCTCATCCT
T7-SSRP1b.fwd	TAATACGACTCACTATAGGGAGGAGATCGGTCGAAAAATTACCGGACC
T7-SSRP1b.rev	TAATACGACTCACTATAGGGAGGAGAGCCAGATAGGCATCTGGT
T7-Spt16a.fwd	TAATACGACTCACTATAGGGAGGAACAAGGGTAATCCCAAGCTGAAGG
T7-Spt16a.rev	TAATACGACTCACTATAGGGAGGAGGTCATGGTTTCCACCTTCT
T7-Spt16b.fwd	TAATACGACTCACTATAGGGAGGAGCGCAACGCGGTGCTGGAGT
T7-Spt16b.rev	TAATACGACTCACTATAGGGAGGAGAGGATAGAGTCCGCCTTCGT

2.2.2.2 T7-DNA template amplification

PCR reactions were carried out using 1X Amplitaq buffer (source, cat #), 200 µM each dNTP, 0.75 µM each of the forward and reverse primers, 250 ng cDNA and 25 units of Amplitaq (Roche Applied Science, 11435094001) in a total volume of 500 µl. The PCR mixtures were aliquoted into eight 0.25ml PCR tubes and amplified using an Eppendorf thermocycler. The following conditions were used for PCR amplification: initial denaturation at 94°C for3 min, followed by 35 cycles of 94°C for 30s, 55°C for 30s and 72°C for 1 min and a 10 min final extension at 72°C. Reaction products were stored at 4°C until use.

Following PCR amplification, the eight reactions were pooled, and the amplicons were purified using the QiaQuick PCR purification kit (Qiagen #28104). Amplicons were eluted from the column into 50 µl of TE (10 mM Tris HCl pH 7.5, 1 mM EDTA). DNA concentration of the final products were measured using a NanoDrop2000c Spectrophotometer, and the quality of the amplicons assessed by electrophoresis on a 2% agarose gel in 0.5 X TBE.

2.2.2.3 DsRNA synthesis

250 µl transcription reactions were assembled at room temperature and contained 1X of 5X HY buffer (1M HEPES-KOH pH 7.5, 100mM MgCl2, 200mM DTT, 10mM Spermidine-HCl and 0.5mg/ml RNAse/DNase free BSA in water), 5mM each of ATP, CTP, GTP, and UTP, 200 U of RNasin, 2.5 U Pyrophosphatase, 5µg of the template DNA and 150 Units T7 RNA polymerase (Invitrogen #18033-019) in a reaction. Reaction mixtures were incubated overnight at 37°C. The template DNA was removed by incubating with 5 U of DNAse RQ1 (Promega #M6101) at 37°C for 30 minutes. The newly synthesized RNA was purified using the QIAGEN RNAeasy midi kit (Qiagen #75144) and eluted in 500 µl of water. The RNA concentration was measured using the NanoDrop2000c Spectrophotometer and adjusted to 1 mg/ml, 10 mM Tris pH 7.5, and 100 mM NaCl. To anneal RNAs, samples were heated for 5 min at 95^{o}C in a heat block in a dry bath incubator. The heat block was removed from the incubator and allowed to cool for 1 hour on the benchtop at room temperature. The newly made dsRNA was run on a 2% agarose gel to check the quality.

2.2.3 DsRNA-mediated knockdown of FACT

To deplete FACT subunits, dsRNA-mediated knockdowns were performed in Drosophila embryonic S2 cells using either one of the two non-overlapping double stranded RNAs targeting either SSRP1 or Spt16. Double stranded eGFP was used as a non-targeting control to rule out off-target effects. All experiments were performed in duplicate. For each replica, about 2×10^6 cells were seeded in each well of a 6-well dish (35 mm) and incubated at 25°C for 30 min to allow cells to attach to the bottom of the plate. Cells were then treated by adding 10 µg dsRNA to each well and incubated for 48 hours at 25 degrees Celsius. A second dose of 10 µg dsRNA/well was added and the cells harvested after 48 hours to study the loss of function phenotypes. Cells were washed twice using 1 ml 1X PBS by centrifugation at 1500 x g for 5min at 4°C and resuspended in 500 µL 1X PBS. A western blot was performed using 10% of the cells to check the knockdown specificity and efficiency. A knockdown of at least 80% was considered acceptable for the downstream experiments.

2.2.4 Custom made SSRP1 and Spt16 antibodies

2.2.4.1 Cloning of full length SSRP1 and Spt16/dre4 constructs

Anti-SSRP1 and anti-Spt16/dre4 polyclonal antibodies were generated by independently injecting rabbits (YenZym Antibodies, LLC) with full length SSRP1 and Spt16/dre4 recombinant proteins bearing an N –terminal Pseudomonas endotoxin. Antibody against SSRP1 and Spt16/dre4 were then affinity purified from sera using CNBr column previously coupled with his-tag purified SSRP1 and Spt16/dre4. Full length SSRP1 and Spt16/dre4 were cloned into Gateway pDONR™221 vectors yielding entry clones of SSRP1 and Spt16/dre4. Briefly, attB-modified custom primers were used to generate attB-

PCR products of SSRP1 and Spt16/dre4 using the Drosophila cDNA as template. To clone attB-PCR products into the pDONR221 donor vector, 50 fmol of attB-SSRP1 PCR product and 50 fmol of attB-Spt16/dre4 were independently mixed with 150 ng pDONR™221 plasmid vector in 1 X TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0), and 2 µl BP Clonase™ in a 10 µl reaction and incubated at 25 °C for 1 h. To stop the reaction, 2 µg Proteinase K solution were added to the BP reaction and incubated at 37°C for 10 min. Competent E. coli cells were then transformed using 2 µl of the cloned products and selected for Kanamycin resistance. M13 primers and customized sequencing primers were used to screen for positive clones. Next, positive clones of pDONR221-SSRP1 and pDONR221-Spt16/dre4 entry clones were each sub-cloned into Gateway destination vectors pVCH6 (Vector has N-terminal Pseudomonas endotoxin and C-terminal 6X His tag) and PET42 (Vector has C-terminal 6X His and V5 epitope) by performing the LR recombination reaction. Again, competent E. coli cells were transformed using 2 µl of the cloned products and selected for ampicillin resistance. M13 primers and customized sequencing primers were used to screen for positive clones.

Table 2-2 Primers for cloning FACT into Gateway vectors

Primer name	5' - 3' sequence
*attB*2-SSRP1.rev	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTA ATC ACT GGC CTC ATC TTC
*attB*1-SSRP1.fwd	GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT AAT GAC AGA CTC TCT GGA G
*attB*2-Dre4.rev	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTA ATG TCG CGA CTT CTT CGA
*att*B1-Dre4.fwd	GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT AAT GTC GAG CTT TGT GCT G
SqP SSRP1_seqR_06	CTG CTC GGT CAT CTT TAG GC
SqP SSRP1_seqR_05	AGA TGG TCT TCG ACT CCT TG
SqP SSRP1_seqR_04	GTC TTG CCG TGC AGC TGG AA
SqP SSRP1_seqR_03	GAT AAA GTT ACC GGG TCC GG
SqP SSRP1_seqR_02	CTT GCC CAT GTT GCT GAC AT
SqP SSRP1_seqR_01	TCT TTC TTC TCG GAC TTC TT
SqP SSRP1_seqF_3	GAG CTG AAG GAC AAG TCC AA
SqP SSRP1_seqF_2	GCT CAA GAA CGG AAC TGT TC
SqP SSRP1_seqF_1	ATG AGC AAG GCC TCG GTC ATC TCG
SqP D221-Dre4.rev_05	TTG CAG GGC TTT GAT GAG CT
SqP D221-Dre4.rev_04	CGT AGA CGT CGC ACA GCT TT
SqP D221-Dre4.rev_03	AGC TCC TTG ACT TCC GGC TC
SqP D221-Dre4.rev_02	GAA ATG CAG AAG AAT GAT CA
SqP D221-Dre4.rev_01	TCC AGC CGC CTT GTT CAA AG
SqP D221-Dre4.fwd_05	TGG CTC TGA TGA GGA ATC TG
SqP D221-Dre4.fwd_04	AGA CTG CCT TCA AGA GCT TC
SqP D221-Dre4.fwd_03	GCC ACA ATG GGT CGC AAC GA
SqP D221-Dre4.fwd_02	GGG CAT CTC GAA TCT GAC CA
SqP D221-Dre4.fwd_01	ACA ACA TAC GCA AGG CAT CG

2.2.4.2 Expression of FACT constructs in Bacteria

Positive clones of pDONR™221-SSRP1 and pDONR™221-Spt16/dre4 were sub cloned into pET-DEST42 and pVCH6 gateway destination vectors. BL21 (DE3) Rosetta pLysS (Novagen) were independently transformed with pET-DEST42-Spt16/dre4, pET-DEST42-SSRP1, pVCH6-SSRP1, and pVCH6-Sp16/dre4 plasmids, plated on LB plates, and incubated overnight at 37 °C. A colony from each plate was picked and used to inoculate 150 ml of 1 X LB supplemented with 34 µg/ml chloramphenicol and 50 µg/ml ampicillin,

and incubated at 37˚C overnight on a shaker. For each construct, Six 1 L 1X LB supplemented with 34 µg/ml chloramphenicol and 50 µg/ml ampicillin were inoculated with 25 ml of the overnight culture and incubated at 37°C on the shaker until culture reached O.D.600 = 0.7 AU. Prior to induction, 1 ml of the culture were taken, spun down, and resuspended in 140 µl 1X sample buffer in 0.1 M DTT for western blot analyses. The recombinant protein expression were then induced by adding 1 ml of 1 M IPTG to the culture at O.D.600 0.7 AU and incubated for 3 hours with agitation at 37˚C. Another 1 ml aliquots were taken for western blot analyses. The remaining bacteria cultures were harvested in 1 L bottles by centrifuging at 6,000 x g for 15 min at 4˚C. Cell pellets were washed once using 50 ml 1X PBS, centrifuged at 6,000 x g for 15 min at 4˚C, and stored at -80°C

2.2.4.3 Cell lysis and fractionation of pVCH6-SSRP1 and pVCH6-Spt16/dre4 antigens under denaturing conditions

PVCH6-SSRP1 and pVCH6-Spt16/dre4 recombinant proteins were purified under denaturing conditions. Prior to purification, bacteria pellets harvested from 2 L cultures were each resuspended in ~ 150 ml of lysis buffer (50 mM Tris-Cl pH 7.5, 1 M NaCl, 20 mM Imidazole pH 7.5, 6 M Urea, 0.1 % Triton X-100, 1x EDTA free Protease Inhibitor, Roche), and sonicated using the Branson digital sonifier micro tip at 50 % amplitude for 5 cycles, where each cycle is 1 min burst of sonication followed by 60 s pause on ice. Soluble proteins were then recovered from the supernatant following centrifugation at 45,000 RPM for 45 min at 4˚C. The supernatant were filtered once in a 0.45 µm CN filter followed by a 0.20 µm CN filter.

2.2.4.4 Nickel affinity purification of pVCH6 antigens under denaturing conditions

Purification of pVCH6 proteins were performed under denaturing conditions on the FPLC system using a Ni2+-charged 1 ml HiTrap Chelating HP column. Using a syringe, the 1 ml HiTrap Chelating HP column was washed with 5 column volumes of water, loaded with 2 ml 0.5 M NiSO4, and washed twice with 5 column volumes of water (1 drop/s). Buffer A (50 mM Tris-Cl pH 7.5, 1M NaCl, 6M Urea, 0.1% Triton X-100) and buffer B (50 mM Tris-Cl pH 7.5, 1M NaCl, 500 mM Imidazole, 6M Urea, 0.1% Triton X-100) were installed and a "pump wash" was run at 4°C on the AKTA purifier 10 (GE Healthcare). Next, the Ni2+-charged 1 ml HiTrap Chelating HP column was installed, the samples were loaded on to the injection inlet, and a 20 column volume gradient from 4 % to 100 % buffer B was run. The flow through were recovered into 50 ml conical tubes from valve 3 outlet, and 500 µl eluate fraction were collected during the gradient phase into capless tubes using the fraction collector. Protein gels of the un-induced culture, induced culture, purification flow through, and the eluate fractions were run and stained using Coomasie blue. The peak fractions were collected and pooled together

2.2.4.5 Cell lysis and fractionation of pET-DEST42 recombinant proteins under native conditions

Prior to purification of pET-DEST42 recombinant proteins, pellets were resuspended in 50 ml lysis buffer (50 mM Tris-Cl pH 7.5, 1 M NaCl, 20 mM, 2mM MgCl2Imidazole pH 7.5, 0.1% Triton X-100, 1 x Protease Inhibitor Roche EDTA free), and finally treated with 1250 U Benzonase I. DNA was sheared again by sonicating lysates on ice using the Branson digital sonifier micro tip at 50 % amplitude for 5 cycles (where each cycle is 1 min burst of sonication followed by 60 s pause). Soluble proteins were

recovered in the supernatant following centrifugation at 45,000 RPM for 45 min at 4°C. The supernatant were filtered once in a 0.45 µm CN filter followed by a 0.20 µm CN filter.

2.2.4.6 Nickel affinity purification of pET-DEST42 recombinant proteins under native conditions

Purifications of pET-DEST recombinant proteins were performed under native conditions on FPLC System using a Ni2+-charged 1 ml HiTrap Chelating HP column. Using a syringe, the 1ml HiTrap Chelating HP column was washed with 5 column volumes of water, loaded with 2 ml 0.5 M NiSO4, and washed twice with 5 column volumes of water (1 drop/s). Buffer A (50 mM Tris-Cl pH 7.5, 1M NaCl) and buffer B (50 mM Tris-Cl pH 7.5, 1M NaCl, 500 mM Imidazole) were installed, and a "pump wash" was run at 4°C on the AKTA purifier 10 (GE Healthcare). Next, the Ni2+-charged 1 ml HiTrap Chelating HP column was installed, the samples were loaded on to the injection inlet, and a 20 column volume gradient from 4 % to 100 % buffer B was run. The flow through were recovered into 50 ml conical tubes from valve 3 outlet, and 500 µl eluate fraction were collected during the gradient phase into capless tubes using the fraction collector. Protein gels of the uninduced culture, induced culture, purification flow through, and the eluate fractions were run and stained using Coomasie blue. The peak fractions were collected and pooled together.

2.2.4.7 Antigen coupling to Cyanogen bromide-activated-Sepharose 4B CNBr-Sepharose beads

Purified recombinant proteins from the Ni-column were dialyzed in 500 ml coupling buffer (50 mM Na-Borate, pH 8.5, 500 mM NaCl, and 2 M urea) for 3 hours (buffer was replaced every hour with fresh buffer). 50 µl of the dialyzed input proteins

were saved to be used to evaluate the coupling efficiency. 0.5 g CNBr-Sepharose 4B (GE Healthcare 17-0430-1) were swelled in 50 ml of cold 100 mM HCl and incubated on the rocker for 30 min at 4 ˚C. The swollen CNBr-Sepharose beads were washed on a fritted funnel using 200 ml of cold 100 mM HCl under very light vacuum. The washed beads were transferred into a 15 ml conical tube, and supernatant discarded after centrifugation at 750 x g for 30 s. Next, beads were equilibrated in 10 ml of cold coupling buffer and the supernatant discarded after centrifugation at 750 x g for 30 s. 2 ml 50% equilibrated CNBr-bead slurry were mixed with 2.5-8 mg of the dialyzed recombinant protein and incubated on the rocker at room temperature for 2-4 h. the ligand-bound CNBr-beads were then spun at 750 x g for 30 s and 100 µl of the supernatant were saved for evaluating the coupling efficiency. The remaining CNBr-Sepharose reactive groups were blocked by adding 1 ml of 1M Ethanolamine pH 8.0 and incubated at room temperature for 1 h on the rocker. Next, ligand-bound beads were spun for 2 min at 750 x g, and the supernatant removed. At this stage, the beads were transferred to columns and washed thrice using 5 ml buffer A (50 mM Na-Acetate pH 4.0, 500 mM NaCl), thrice using 5 ml buffer B (50 mM Na-Borate pH 8.0, 500 mM NaCl), and twice using 5 ml 1 X PBS. 60 µl of 2% sodium azide were added to columns prior to storage at 4 ˚C.

2.2.4.8 Affinity purification of SSRP1 and Spt16 antibodies from Serum

1 ml antigen-bound beads (1ml) were mixed with serum in a 50 ml falcon tube and incubated at 4˚C overnight on a rotator. Sera/beads mixture were loaded onto a 10 ml BioRad econo-column fitted with a 16G1 needle connected to 20-30 cm of 0.062”/0.082” I.D./O.D. polyethylene. The columns were then rinsed five times using 10 ml 1 X PBS. Columns were washed again sequentially using, 10 ml of Wash A (50 mM Tris pH 8.0, 120 mM NaCl, 0.5 % NP-40), 10 ml of Wash B (50 mM Tris pH 8.0, 1 M LiCl, 0.5 % NP-40), 10

ml of Wash A and 10 ml of 1 X PBS. After the washes were completed, the column was connected to a fraction collector to collect 400-500 µl eluate into capless Eppendorf containing 7.5 µl of 2 M Tris-base. Antibodies were eluted by adding 5 ml of elution buffer (150 mM NaCl, 50 mM Glycine-Cl pH 2.5) to column. Bradford assay were performed by mixing 5 µl of each fraction to 100 µl of Bradford reagent, Coomassie Plus (Thermo SCIENTIFIC, 23238). Fractions with high concentration of antibody (measured using Bradford assay) were pooled together. For storage, 2% Sodium azide was added to each pooled fraction to 0.02 % final concentration

2.2.5 Western blotting

Saved aliquots of cell suspension were resuspended in NuPAGE LDS sample buffer (Invitrogen, NP0007) containing 0.1 M DTT, sonicated using a BRANSON digital sonifier at 4°C using 100 % amplitude for 3 cycles (where each cycle is a 60 s burst of sonication, followed by a 60 s pause), and heated at 70 °C for 10 min. 10 µl aliquots of reduced samples were loaded into wells of a NuPAGE Novex 4-12 % bis-tris SDS-PAGE gradient gel (Life tech, NP0322BOX) and run using 1 X MES (50 mM MES, 50 mM Tris Base, 0.1 % SDS, 1 mM EDTA, pH 7.3) buffer and 500 µl NuPAGE antioxidant (Invitrogen, NP0005) at 120 v for ~2 h. The resolved proteins were then transferred onto a pre-activated Immobilon-FL PVDF transfer membrane (EMD Millipore, IPFL00010) using a Genie electrophoretic transfer system (Idea Scientific Co.). The transfer was carried out using 1 X transfer buffer (25 mM Tris base, 188 mM Glycine, 10 % Methanol) at 24 v for 1h 15 min. After transfer, membranes were washed once with 1X PBS and blocked for an hour using Odyssey blocking buffer (LI-COR, 927-40000). To probe, the membranes were incubated in 1:5000 primary antibodies in Odyssey blocking buffer at room temperature for at least an hour. Unbound primary antibodies were washed 3 x 5 min, by incubating membrane in buffer

containing PBST (1 X PBS + 0.1 % Tween-20) at room temperature with gentle shaking. In order to visualize, blots were incubated the blot in a 1:10000 fluorescently-labeled secondary antibody in Odyssey blocking buffer for 45 min, washed, and scanned using the Odyssey fluorescence imaging system (LI-COR Biosciences).

2.2.6 Chromatin Immunoprecipitation of FACT subunits

2.2.6.1 Formaldehyde crosslinking, sonication and immunoprecipitation

Chromatin immunoprecipitation was performed as described (Lee et al., 2006). Briefly, Formaldehyde crosslinking was performed by adding an equal volume of 2% formaldehyde solution (900 µL 16% formaldehyde in 7.3 ml 1 X PBS) to the resuspended cells and incubated at room temperature for 20 min. The formaldehyde crosslinking was quenched by adding 1/20 volume of 2.5 M glycine to the cells in suspension. The cells were then washed twice using ice-cold 0.125 M glycine in 1 X PBS. Next, the cell pellet was resuspended in 500 µL buffer A2 (15mM HEPES pH 7.5, 140 mM NaCl, 1mM EDTA, 0.5 mM EGTA, 1 % Triton-X 100, 0.1% Sodium deoxycholate, 1 % SDS, 0.5 % N-lauroylsarcosine, protease inhibitor Cocktail-Sigma P8849). The nuclei were pelleted by spinning at 3500 x g for 5min at 4 °C. The supernatant was carefully discarded and nuclei resuspended in 800 µl Buffer A2. The extracted nuclei were sonicated using the Misonix 3000 sonicator at 4°C using output power 3 (9 w power) for 10 cycles (where each cycle is a 10 s burst of sonication, followed by a 60 s pause) and spun at 14000 rpm at 4°C for 10 min. A 30 µL aliquot of the sheared chromatin (supernatant) was saved as input sample to normalize the eluate signals and to check the sonication efficiency. Prior to the immunoprecipitation, 25 µl Pierce Protein A/G Magnetic Beads (Invitrogen, 88803) were washed twice using BSA solution (0.5% BSA in 1X PBS). Next, 10 µg of anti-SSRP1 (T. Tettey and M. Blanchette), 10 µg of anti-Spt16 (T. Tettey and M. Blanchette) and 10 µg

Normal Rabbit IgG (Santa Cruz, sc-2027) in BSA solution were coupled to the washed Protein A/G beads for 4 hours at 4°C. The antibody-conjugated Dynabeads were then washed twice using 0.5% BSA in 1X PBS. The sheared chromatin was then added to the antibody-conjugated Protein A/G beads and incubated overnight at 4°C. The Protein A/G beads-antibody-protein complex were washed twice using a low salt buffer (20mM Tris-HCl, 150mM NaCl, 2mM EDTA, 1% Triton X-100), once with a high-salt buffer (20mM Tris-HCl, 500 mM NaCl, 2mM EDTA, 1% Triton X-100) and once with a LiCl buffer (20 mM Tris-HCl, 250 mM LiCl, 1 mM, 1% Triton X-100, 0.1% NP-40, 0.5% sodium deoxycholate) at 4°C. Washed Protein A/G-antibody-protein complexes were resuspended in 150 µl elution buffer (50mM Tris pH 8.0, 10mM EDTA, 1% SDS), 150 µl TE (10 mM Tris, 1mM EDTA), and 4 µl of 10 µg/µl RNaseA (Sigma; Cat. no. R6513). In parallel, 120 µl elution buffer, 150 µl TE, and 4 µl of 10 µg/µl RNaseA was added to 30 µl of each input sample. Input samples and Protein A/G-antibody-protein complexes were then incubated at 37°C while shaking on the Thermomixer for 30 min. In order to digest the proteins and reverse the crosslinking, 2 µl of Proteinase K solution (Invitrogen; Cat. no. 25530-049) was added to the eluate and incubated overnight at 65°C while shaking on the thermomixer.

2.2.6.2 DNA purification

DNA was extracted from the reverse crossed-linked samples using Phenol:Chloroform:Isoamyl Alcohol. Briefly, an equal volume of Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (Sigma; Cat no. P3803) was added to the samples, vortexed and centrifuged at 15000 x g. The supernatant was carefully collected into fresh tubes containing 12 µl of 5M NaCl (final 200 mM) and 30 µg glycogen (Roche; Cat no. 901 393). DNA was precipitated by mixing the supernatant with 750 µl iced cold 100% ethanol, incubated at -80 °C for 1 h, and spun at 20000 x g for 30min at 4°C. DNA

pellets were washed once using 1ml 70% ethanol, air-dried for 30min and resuspended in nuclease free water.

2.2.6.3 Library preparation

The Molecular Biology Core of the Stowers Institute prepared libraries from the ChIPed DNA using the KAPA HTP kit (Kapa Biosystems), sized selected using the SPRIselect double side size selection kit (Beckman Coulter) and sequenced on the Illumina HiSeq platform.

2.2.6.4 ChIP-seq analysis

This and all bioinformatic analyses described in this thesis were performed with Xin Gao, Alex Chitsazan, Benjamin Story, and Marco Blanchette. The ~30 million reads from each experiment were aligned using bowtie 2 to the Drosophila dm6 genome. The two independent biological replicates (NS1 and NS2) of the eGFP dsRNA treated cells had very good correlation (r = 0.98) and so data generated were averaged. ChIP-seq data generated from the four independent biological replicates of SSRP1 knockdown (derived from biological duplicates of the two unique SSRP1-dsRNA) were averaged since they had very good correlation (r=0.9). Therefore, all the heatmaps and metagene analyses represent averaged values.

2.2.7 RNA-seq

2.2.7.1 FACT knockdown, RNA purification and library preparation

Two non-overlapping, ~ 400 base, dsRNAs were designed to target Spt16 and SSRP1. In addition, dsRNA of eGFP was used as a non-targeting control. All the double stranded RNA mediated knockdown experiments were performed in duplicate.

For each knockdown experiment 2×10^6 cells were seeded per well of a six-well plate and incubated at 25°C for 30 min to allow cells to attach to the bottom of the plate. Cells were treated with 10 µg dsRNA and incubated for 48 hours. The cells were again treated with 10 µg dsRNA and harvested after 48 hours to study the loss of function phenotypes. After 48 hours, cells were transferred into 2ml tubes, washed once with 1ml PBS, and resuspended in 100 µl 1X PBS. About 10% of the cell suspension of each sample was used in a western blot to determine the extent of knockdown. From the remainder of each sample, total RNA was prepared using the QIAGEN RNeasy mini kit (Qiagen, cat# 74106) following the manufacturer's protocol. The purified RNA was eluted in 80 µl of water. The quantity and integrity of the isolated RNA were assessed using the Agilent Bioanalyzer.

The Molecular Biology Core of the Stowers Institute for Medical Research prepared the mRNA libraries following the Illumina TruSeq Stranded mRNA sample preparation protocol (Illumina, 15031047) using the TruSeq Standard mRNA LT Sample Prep Kit, 48; Set A (Illumina, RS-122-2101). Briefly, the polyA+ RNA was purified using oligo dT magnetic beads. The purified mRNAs were fragmented, randomly primed and reverse transcribed to generate first and second strand cDNAs. The 3'-ends of the cDNAs were adenylated, adaptors ligated to the ends and the final cDNA enriched through PCR. The libraries were then sized selected using the Pippen (Sage Science) and validated using Agilent 2100 Bioanalyzer. The barcoded libraries were pooled together and loaded onto a flow-cell for sequencing to generate 50 bp single reads.

2.2.7.2 RNA seq analysis

Reads from two biological replicates for each sample were aligned to the fly genome dm6 using the default parameters of the STAR aligner. Cuffdiff v1.3.0 was used to

quantify RPKM values, to perform differential expression analysis at FDR < 0.05, and to assess statistically sufficient read coverage for each gene.

2.3 Results

2.3.1 DsRNA mediated knockdown of SSRP1 or Spt16 in S2 cells were successful

In order to determine how FACT regulates gene expression globally in higher eukaryotes, levels of FACT subunits were depleted in an RNAi experiment in Drosophila embryonic cell line S2 followed by sequencing of messenger RNA. For every target protein, two non-overlapping, ~ 400 base dsRNAs were designed. In addition, double stranded eGFP RNA was used as a non-targeting control. DsRNA mediated knockdown of SSRP1 or Spt16 led to a significant reduction of the protein level of both FACT subunits as shown in the western blot (Figure 1A). The two unique non-overlapping dsRNA targeting each FACT subunit yielded very similar results, indicating that the observed were not because of “off-target “effect of the dsRNAs. The observation that knocking down either SSRP1 or Spt16 reduced expression of both proteins, is consistent with prior results indicating that depletion of one FACT subunit leads to reduced stability of the other in human cells (Safina et al., 2013).

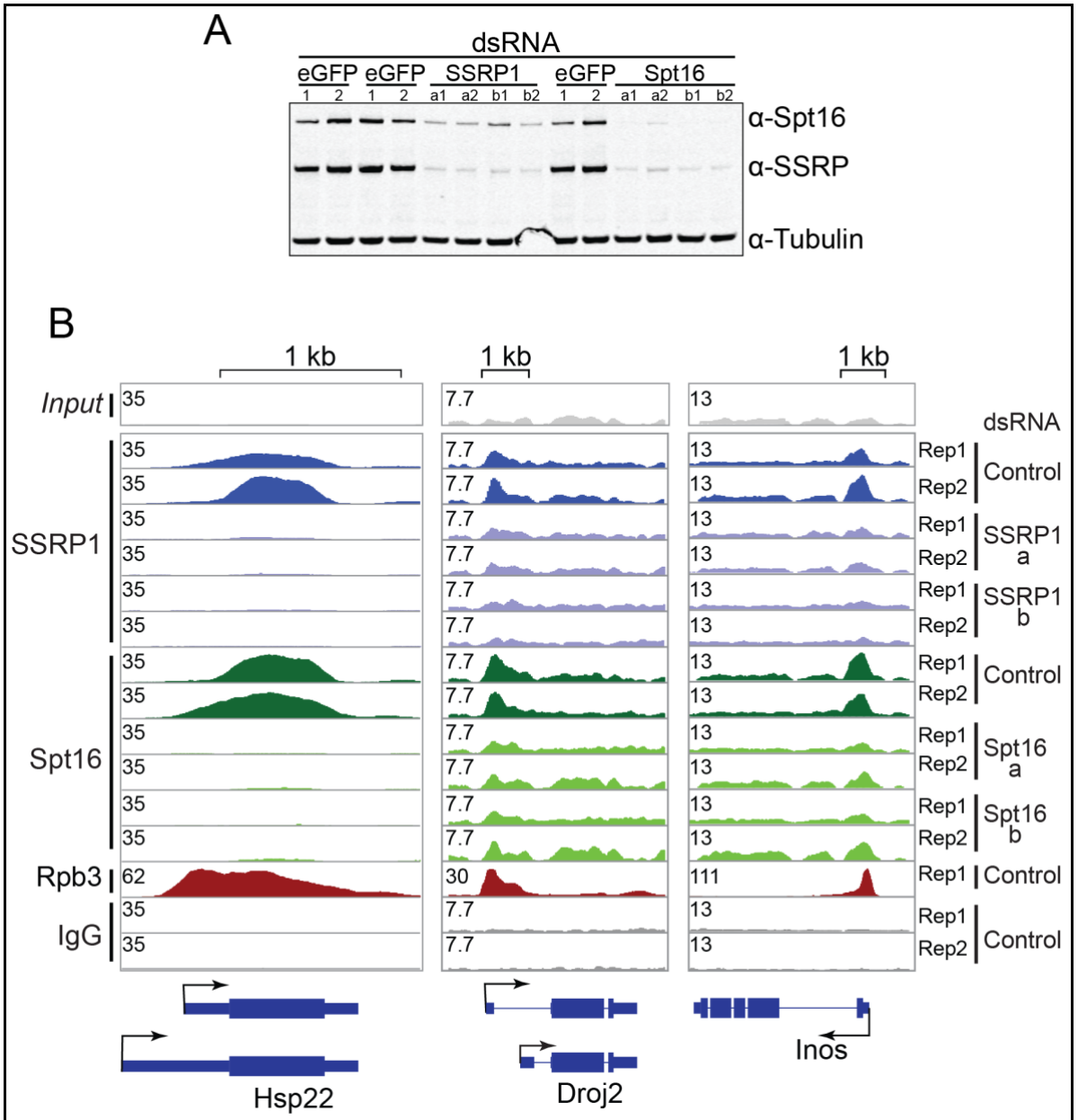

Figure 1. DsRNA mediated knockdown of FACT, and ChIP of FACT were successful. (A) Western blot of Spt16 and SSRP1 in Drosophila S2 cells previously treated with eGFP, SSRP1, or Spt16 double stranded RNAs. Tubulin was used as a loading control. (B) IGV browser snapshot showing Spt16 (colored green) and SSRP1 (colored blue) localization in relation to RNA Pol II (Rpb3 subunit) within the Hsp22, Droj2, and Inos loci. Levels of Spt16 and SSRP1 were reduced in the Spt16 and SSRP1 deficient cells respectively

2.3.2 FACT binds along transcribed regions with bias towards 5' ends

To begin to characterize functions of FACT in higher eukaryotes genome-wide, I generated antibodies against Drosophila FACT subunits SSRP1 and Spt16 and used them to perform chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) experiments in Drosophila S2 cells. To control for specificity of the SSRP1 and Spt16 antibodies, SSRP1 or Spt16 were each depleted using two unique double stranded RNAs (dsRNA). EGFP double stranded RNA was used as a non-targeting control.

The peak finding algorithm, MACS2 (Model-based Analysis of ChIP-seq), was used to identify peaks (MACS 2 $q \leq 1 \times 10\text{-}4$) that occur in all replicate samples. Comparison of peaks identified by ChIP of SSRP1 and Spt16 revealed that the localization of the two proteins was highly correlated, as expected since they function as subunits of the same complex (Figures 1B, and 2B).

At genes showing the greatest enrichment for FACT subunits, exemplified in Figure 1B by Hsp22, SSRP1 and Spt16 ChIP signals were almost completely lost after knockdown of SSRP1 or Spt16, respectively. At other genes, such as Droj2 or Inos, where SSRP1 and Spt16 were enriched to a lesser extent, the ChIP signals were reduced, most clearly near gene 5' ends, but not lost (Figure 1B), perhaps because of the presence of residual SSRP1 and Spt16.

We then partitioned the genome into promoter and early transcribed region (200 bp upstream to 1 kb downstream the TSS), gene body, and intergenic regions as shown in Figure 2A (first panel from left). We observed that more than 80% of the Spt16 peaks and 90% of the SSRP1 peaks fell within promoters and gene bodies, with the majority of these falling within promoter and early transcribed regions (Figure 2A, panels 2 and 3 from left). This observation suggests that subunits of FACT bind along transcribed regions of genes

with a bias towards promoter regions (Figure 2A), consistent with earlier findings suggesting that yeast FACT binds to transcribed regions of specific genes.

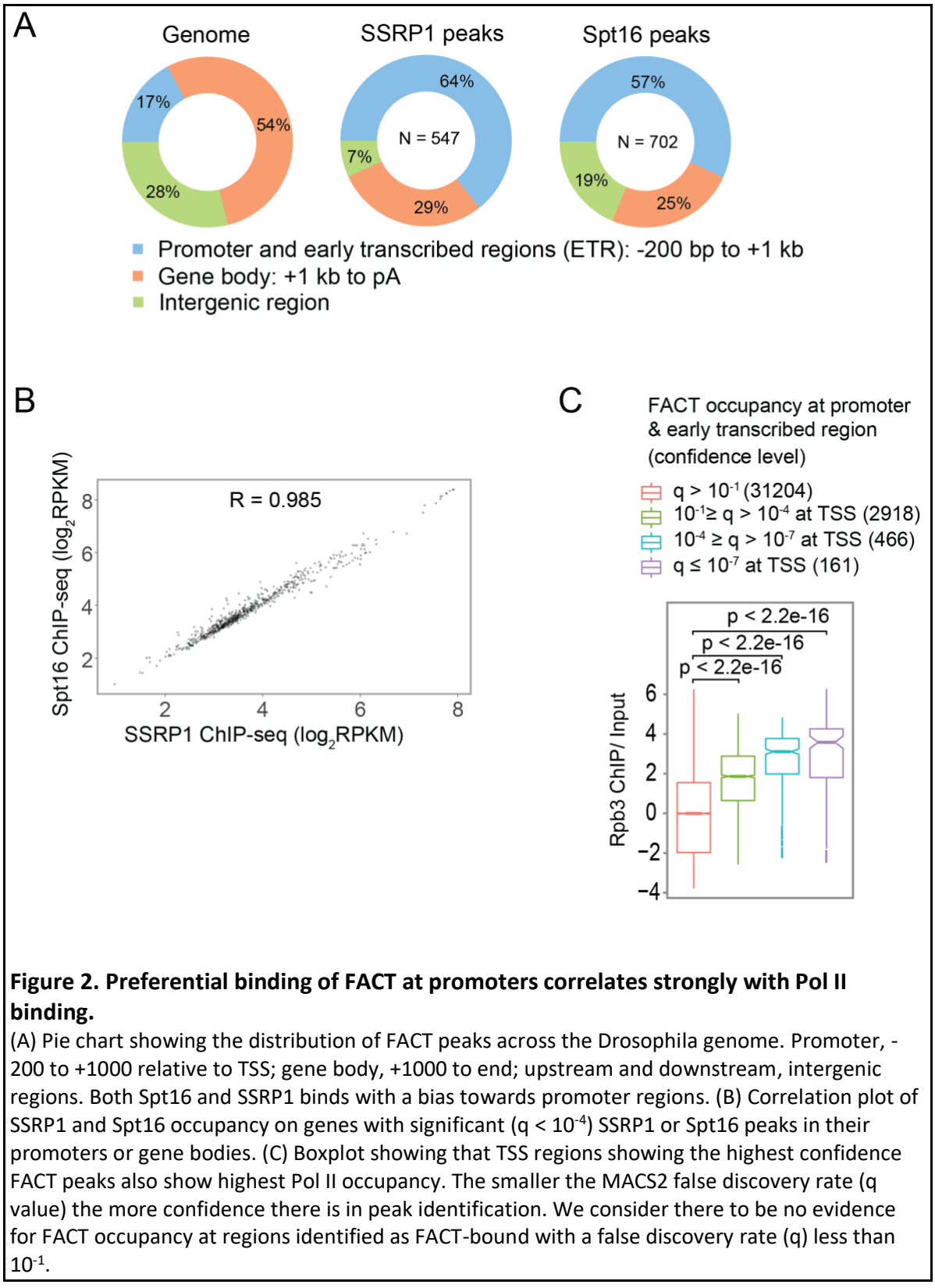

Figure 2. Preferential binding of FACT at promoters correlates strongly with Pol II binding.
(A) Pie chart showing the distribution of FACT peaks across the Drosophila genome. Promoter, -200 to +1000 relative to TSS; gene body, +1000 to end; upstream and downstream, intergenic regions. Both Spt16 and SSRP1 binds with a bias towards promoter regions. (B) Correlation plot of SSRP1 and Spt16 occupancy on genes with significant ($q < 10^{-4}$) SSRP1 or Spt16 peaks in their promoters or gene bodies. (C) Boxplot showing that TSS regions showing the highest confidence FACT peaks also show highest Pol II occupancy. The smaller the MACS2 false discovery rate (q value) the more confidence there is in peak identification. We consider there to be no evidence for FACT occupancy at regions identified as FACT-bound with a false discovery rate (q) less than 10^{-1}.

2.3.3 FACT occupancy positively correlates with Rpb3 occupancy at promoters

To assess the relationship between FACT and Pol II occupancy, we performed a ChIP of total Pol II using an antibody targeting the Pol II subunit Rpb3. As shown in Figure 1B, regions of highest FACT enrichment in the 5' ends of genes overlapped with but tended to be somewhat downstream of promoter-proximally paused Pol II. We then compared Rpb3 enrichment in the 5'-end of genes (200 bp upstream to 1 kb downstream of the TSS) in order to determine whether genes with high confidence FACT peaks in this region tend to be associated with more Pol II. As shown in Figure 2C, gene regions with higher confidence FACT peaks (MACS2 q<10-4) exhibited the highest Pol II occupancy, while regions with no evidence for FACT binding had much lower Pol II occupancy (Figure 2C). Of note, results of MACS2 analysis suggest that FACT might be broadly distributed, as a large number of low confidence peaks (10-1≥q>10-4) were also identified, and these too showed increased Pol II occupancy relative to regions not bound by FACT. These observations are consistent with previous evidence that FACT is associated with transcriptionally active genes in yeast (Mason and Struhl, 2003), colocalizes with Pol II on Drosophila polytene chromosomes (Saunders et al., 2003), and is recruited via HP1c to actively transcribing Pol II in Drosophila (Kwon et al., 2010).

2.3.4 RNAi mediated knockdown of SSRP1 and Spt16 in S2 cells affect RNA transcript abundance significantly

The observation that the degree of FACT binding correlates with RNA polymerase II enrichment around promoters led us to ask whether loss of FACT affects transcriptional outcomes. To address that, I performed a poly-A selected RNA-seq on purified RNA from FACT depleted and control (eGFP dsRNA treated) Drosophila S2 cells. Using an absolute fold change of at least 1.5 and a false discovery rate of at most 0.05, we found that

expression of about 19% of the 7054 expressed genes (genes with CPM ≥ 5 in at least 2 samples) was altered by knocking down either SSRP1 or Spt16. In addition, a good correlation was observed between the effect of knocking down SSRP1 and Spt16 on global gene expression (Figure 3A and B) as would be expected if the two subunits work together in a complex. Overall, we identified about 1298 and 262 genes whose expressions were significantly ($p \leq 0.05$) affected by at least 1.5-fold following SSRP1 and Spt16 depletion respectively. Majority of genes were up-regulated following depletion of either Spt16 or SSRP1 subunit. RNAi mediated knockdown of SSRP1 led to about 68% of the affected genes being up-regulated and 32% downregulated. Similarly, about 87% of the affected genes were up-regulated while 13% down-regulated following knockdown of Spt16. Since FACT was discovered as a facilitator of chromatin transcription (Orphanides et al., 1998) it was also expected that the expression of some genes might be impaired following FACT loss. Our findings, however, suggest that the loss of FACT *in vivo* leads not only to downregulation of genes but also to more gene activation events, which could be expected if reassembly or disassembly of histones over genes is defective, leaving the chromatin in a disordered state conducive to pervasive transcription.

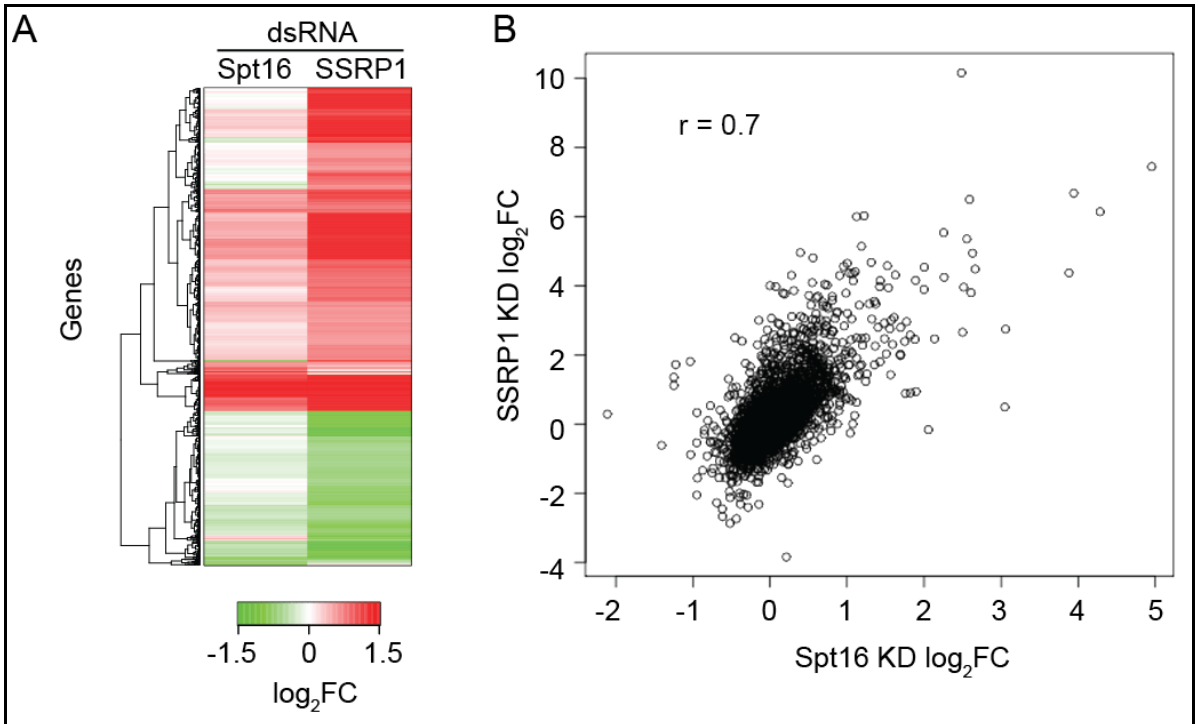

Figure 3. Depletion of FACT subunits alters RNA transcript abundance.
(A) Heatmap showing hierarchical clustering of 1335 genes that change by an absolute fold change of 1.5 (FDR < 0.05) in either SSRP1 or Spt16 knockdown compared to the non-targeting control. Green indicates fold decrease while red denotes fold increase in the FACT dsRNA treated cells compared to the eGFP dsRNA treated cells. (B) Correlation plot showing gene expression changes of all expressed genes following the knockdown of SSRP1 versus Spt16. Y-axis represents fold change after SSRP1 knockdown while X-axis represents fold change after Spt16 knockdown.

2.4 Discussion

The results from this chapter demonstrate that dsRNA-mediated knockdown of FACT subunits successfully depleted SSRP1 and Spt16 proteins. The loss of SSRP1 led to the reduction of the Spt16 subunit and vice versa. This observation suggest that the stability of FACT in Drosophila S2 cells depends on the presence of both SSRP1 and Spt16 proteins consistent with observations made by Safina and colleagues in mammalian cells (Safina et al., 2013).

In addition, using ChIP followed by high-throughput sequencing, I observe that FACT subunits bind preferentially to promoter and early transcribed regions of genes and this correlates with RNA polymerase II enrichment around promoter regions. This

observation supports earlier *in vitro* studies suggesting that FACT regulates transcription initiation.

Furthermore, depletion of FACT led to both up- and down-regulation of many transcripts, consistent with the model that FACT acts as both a negative and positive regulator of gene expression. Although there was a good correlation between the effects of SSRP1 and Spt16 depletion, the magnitude of gene expression changes was consistently greater following SSRP1 depletion.

At the present time, we do not understand the mechanism(s) responsible for the apparent differences in SSRP1 and Spt16 function. As shown in Figure 1A, comparable reductions in SSRP1 levels were observed in cells treated with dsRNAs targeting either SSRP1 or Spt16. In contrast, more Spt16 remains in cells treated with SSRP1 dsRNAs than Spt16 dsRNAs, raising the possibility that free Spt16 may remain after SSRP1 knockdown. Consistent with the possibility that unbalanced expression of Spt16 and SSRP1 could give rise to phenotypic changes, Spt16 was originally identified in an overexpression screen for genes that would suppress the phenotype of δ insertion mutations by changing transcription, and it was proposed that increased levels of Spt16 might alter function(s) of protein with which it normally interacts (Malone et al., 1991).

The finding that SSRP1 mutation or depletion can have stronger phenotypes than Spt16 mutation or depletion is not unprecedented. For example, microarray analyses in human cells showed that expression of a common set of about 120 genes was altered by siRNAs against either SSRP1 or Spt16; however, an additional 12 showed changes in expression only in SSRP1 knockdown cells (Li et al., 2007). While Arabodopsis genetically modified to express reduced levels of SSRP1 are phenotypically generally similar to plants expressing reduced Spt16; however, the SSRP1 mutant has a stronger phenotype when combined with a TFIIS mutant (Antosz et al., 2017). In addition, in a population of plant

progenitor cells, heterozygous SSRP1 mutants, but not Spt16 mutants, exhibit defects in DNA methylation and regulation of parentally imprinted genes (Ikeda et al., 2011).

Chapter 3

Roles of FACT in Histone Modification and Chromatin Architecture

3.1 Introduction

Nucleosomes, the basic unit of chromatin have been shown to be a major impediment to transcription initiation and elongation. Cells have therefore evolved a wide range of enzymes that slides, reorganizes, disassembles, reassembles, ejects, and covalently modify histones in an ATP dependent or independent manner in order for transcription to take place. The FACT complex is one of the highly potent histone chaperones that disassembles and reassembles nucleosomes in an ATP independent manner to stimulate transcription initiation and elongation *in vitro* (Belotserkovskaya et al., 2003). Work from a number of labs have hinted at the possible involvement of FACT in histone modifications and chromatin architecture organization based on studies in yeast on specific genes. Co-immunoprecipitation experiments performed in SETD2-depleted cells suggest H3K36me3 recruits the elongation factor FACT to transcribed regions (Carvalho et al., 2013), leading to the proposal that Set2 regulates recruitment of FACT through H3K36me3. In addition, Stanlie and colleagues observed that reducing levels of FACT leads to a reduction of H3K4me3 levels over Sμ and Sα regions, suggesting that FACT-dependent changes in H3K4me3 plays a role in Ig class switch recombination (Stanlie et al., 2010). A more recent study has demonstrated that yFACT functions to restrict H2A.Z to promoter locations since mislocalisation of H2A.Z into genic locations in FACT mutants leads to increase in cryptic transcription (Jeronimo et al., 2015).

In this chapter, I begin to explore the roles of FACT in histone modifications and chromatin architecture in higher eukaryotes. My previous observations suggest that FACT binds preferentially to promoter regions and depletion of FACT leads to deregulation of

many transcription events *in vivo*. Based on these findings, I hypothesized that transcription is deregulated in FACT knockdown cells because nucleosomes that normally would block accessibility to promoters and transcribed regions are not properly re-assembled in the wake of Pol II transcription. To begin addressing this hypothesis, I performed an MNase-seq and ChIP of H3 and in cells depleted of FACT to measure nucleosome positioning and occupancy respectively. In addition, I performed ChIPs of the transcription related histone variant H2Av and histone modifications including H3K4me3, H3K36me3 and H3K56 ac to assess chromatin state globally.

3.2 Materials and Methods

3.2.1 Micrococcal nuclease assay

3.2.1.1 FACT knockdown, formaldehyde crosslinking and micrococcal nuclease digestion

Prior to the MNase assay, 2 million cells were seeded per well in 6-well dishes and depleted of SSRP1 and Spt16 in an RNAi experiment as described (Clemens et al., 2000). Double stranded RNA made from eGFP (dseGFP RNA) was used as a non-targeting control. The extent of protein knockdown was determined through western blot. The mono-nucleosomal DNA was prepared essentially as described (Gilchrist et al., 2008). Briefly, the treated cells were harvested into 2ml tubes and pelleted at 1000 x g for 2 min. Cells were resuspended in 1 mL 1X PBS, pelleted by centrifugation at 1500 x g for 5min at 4°C, and resuspended in 500 µl 1X PBS. 25 µl aliquots of each cell suspension were removed for western blot analysis. To perform the formaldehyde crosslinking, an equal volume of 2% formaldehyde in 1X PBS was added to cell suspensions and incubated for 45 seconds at room temperature in 2 ml Eppendorf tubes in a Thermomixer with agitation at 1400 rpm. Formaldehyde cross-linking were quenched by adding 50 µl of 2.5 M glycine

and incubating for 5 minutes at room temperature in a Thermomixer with agitation at 1400 rpm. Crosslinked cells were then pelleted at 1000 x g for 2 min at 4 °C and washed using 1X PBS. Cells were permeabilized by resuspending in buffer A (10 mM Tris-HCl pH 8, 10 mM KCl, 3 mM $CaCl_2$, 0.34 M sucrose, 10% glycerol, 1 mM DTT, Protease Inhibitor (Roche complete, EDTA free, lot#13744100), and 0.1% Triton X-100) and incubating on ice for 15 minutes with intermittent up and down pipetting using a P-200. Cell lysates were spun at 2500 x g for 2 min at 4 °C, and the pelleted nuclei were washed once with Buffer B (4 mM EDTA, 0.2 mM EGTA, Protease Inhibitor). Washed nuclei were then resuspended in 300 µl cold Buffer B and incubated on ice for 15 minutes. Chromatin were isolated by centrifugation at 2500 x g, for 2 min at 4 °C and supernatant discarded. Next, the pelleted chromatin were resuspended in Digestion Buffer (15 mM Tris-Cl, pH 8, 60 mM KCl, 15 mM NaCl, 1 mM DTT, 0.25 M sucrose, 1 mM CaCl2). Twenty units of MNase were used to digest 200 µL of chromatin. A time course of micrococcal nuclease digestion (5, 10 and 15 minutes at 25°C) were performed for each sample. Stop Solution containing 20mM Tris pH 7.4, 2% SDS, 25 mM EDTA were added, and samples were incubated for 90 min at 65°C to reverse the cross-linking. To digest proteins, reverse-crosslinked samples were incubated with 1 µl proteinase K (Roche, 03115828001) at 37°C overnight in thermomixer at 1200 rpm. DNA was extracted using ethanol precipitation. The purified DNA was resuspended in 10 µl 1 X TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) buffer and treated with 10 µg RNase A for 15 minutes at 37^{o}C. The purified DNA were resolved on a 1.5 % agarose gel with 0.5 X TBE (40 mM Tris-HCl, 45 mM boric acid, 1 mM EDTA). Mono-nucleosomal DNA from samples digested for a time sufficient to yield ~80 % DNA from mononucleosomes (~150 bp) and ~20% DNA from di- and trinucleosomes (~300 and ~450 bp, respectively) were used for further analysis. The mono-nucleosomal DNA were

excised from the gel, purified using the QIAquick gel extraction kit (Qiagen, cat. no. 28704) and resuspended in 10 µl of buffer EB.

3.2.1.2 MNase-seq library preparation

The Molecular Biology Core of the Stowers Institute for Medical Research prepared libraries from the purified DNA and performed paired end sequencing. Libraries were prepared using the KAPA HTP Library Preparation Kit v3.12 for Illumina® Platforms (KAPA Biosystems, KK8234) on the Sciclone Robot following the manufacturer's protocol. Prior to library preparation, dsDNA was quantified using the QUBIT high sensitivity assay. Briefly, for library preparation 10 ng of the DNA was end repaired, adenylated, indexed, and enriched using PCR. Libraries were then quantified using high sensitivity Qubit and High sensitivity Bioanalyser assays. The barcoded samples were pooled together and run on 4 lanes of the Illumina HiSeq using the rapid run protocol. 100 bp paired end reads of the samples were sequenced.

3.2.1.3 MNase-seq analysis

15 million reads were mapped using Bowtie 2 (parameter –no mixed, - no discordant) to the D. melanogaster genome dm6 Ens84. Average insert size was calculated using Picard CollectInsertSizeMetrics. Reads with lengths between 120 and 180 (representing mono-nucleosomal DNA) were used for downstream analysis. Reads were then resized to 146 bases from the center of each read.

3.2.2 Chromatin Immunoprecipitation of H3K4me3, H3K36me3, H2Av and H3K56ac

3.2.2.1 Cell culture, dsRNA mediated knockdown of FACT and western blot

Prior to Chromatin immunoprecipitation (ChIP) of histone H3, histone marks (H3K36me3, H3K4me3 and H3K56ac), and histone variants (H2Av), I performed an

independent knockdown of Spt16 and SSRP1 in Drosophila embryonic S2 cells as described above.

3.2.2.2 Formaldehyde crosslinking, sonication and immunoprecipitation

Chromatin immunoprecipitation was performed as described above using 10 µg anti-H3 (ab1791, Abcam), 10 µg of anti-H3K36me3 (ab9050, Abcam), 10 µg of anti-H3K4me3 (ab8580, Abcam), 10 µL of rabbit anti-H2AvD serum (Rabbit anti-Drosophila H2A variant, (Leach et al., 2000)) in BSA solution coupled to the washed Dynabeads for 4 hours at 4°C.

3.2.2.3 DNA purification

DNA was extracted from the reverse crossed-linked samples using Phenol:Chloroform:Isoamyl Alcohol resuspended in nuclease free water as described above.

3.2.2.4 Library preparation

The Molecular Biology Core of the Stowers Institute prepared libraries from the ChIPed DNA using the KAPA HTP kit (Kapa Biosystems), sized selected using the SPRIselect double side size selection kit (Beckman Coulter) and sequenced on the Illumina HiSeq platform.

3.2.2.5 ChIP-seq analysis

The ~30million reads from each experiment were aligned to the drosophila dm6 Ens84 genome using default parameters of Bowtie 2. ChIP-seq reads were then extended to 150 bp.

3.3 Results

3.3.1 Loss of FACT is associated with a subtle decrease in nucleosome occupancy and positioning

Based on our results and published data, I hypothesized that "transcription is deregulated in FACT knockdown cells because nucleosomes that normally would block accessibility to promoters and transcribed regions are not properly reassembled in the wake of Pol II transcription". To begin addressing this hypothesis, I mapped nucleosome positions using micrococcal nuclease digestion followed by sequencing (MNase-seq), and measured histone occupancy by performing a ChIP of histone 3 followed by high throughput sequencing.

Metagene analysis of ChIP-seq of endogenous H3 across TSS and TES showed accumulation of H3 across the entire breath of transcribed region. I observed only minimal changes in H3 following the knockdown of FACT subunits compared to the eGFP dsRNA treated S2 cells (Figure 4A). The effects of the loss of either SSRP1 or Spt16 on H3 levels correlate very strongly. Since the depletion of the SSRP1 subunit showed the biggest change in RNA transcript accumulation compared to the loss of the Spt16 subunit of FACT, one would have expected that SSRP1-dependent change in H3 would have been more pronounced. This observation was unexpected, as yeast FACT has previously been demonstrated to regulate histone occupancy. One possible interpretation is that other chaperones or chromatin modulators in S2 cells may compensate for the loss of FACT.

To determine nucleosome positioning, I performed MNase-seq by subjecting cross-linked chromatin isolated from eGFP dsRNA and SSRP1 dsRNA treated S2 cells to micrococcal nuclease (MNase) digestion followed by paired end sequencing of the mono-nucleosomal DNA. Micrococcal nuclease preferentially cuts the linker DNA between

nucleosomes leaving footprints of mainly nucleosome protected DNA (Figure 4B). An average gene plot of the read distribution of mono-nucleosomal DNA around the TSS and transcribed regions of all expressed genes showed a nucleosome depleted region around the TSS and well positioned +1, 2, and 3 nucleosomes within transcribed regions (Figures 4C and 4D). Next, we observed a subtle decrease of MNase-seq signal at the +1 nucleosome in the Spt16 (Figure 4C) and SSRP1 (Figure 4D) depleted cells compared to the eGFP dsRNA treated S2 cells. Overall, however, I find no major changes neither in nucleosome occupancy (measured by chromatin immunoprecipitation of Histone H3), nor in nucleosome positioning (determined by MNase-seq) in the *Drosophila* FACT knockdown compared to previous reports of defects in yeast FACT mutants. One explanation for this discrepancy may be that as transcription programs in Drosophila are more complex than those of yeast, Drosophila S2 cells may have evolved mechanisms that compensate for the loss of FACT.

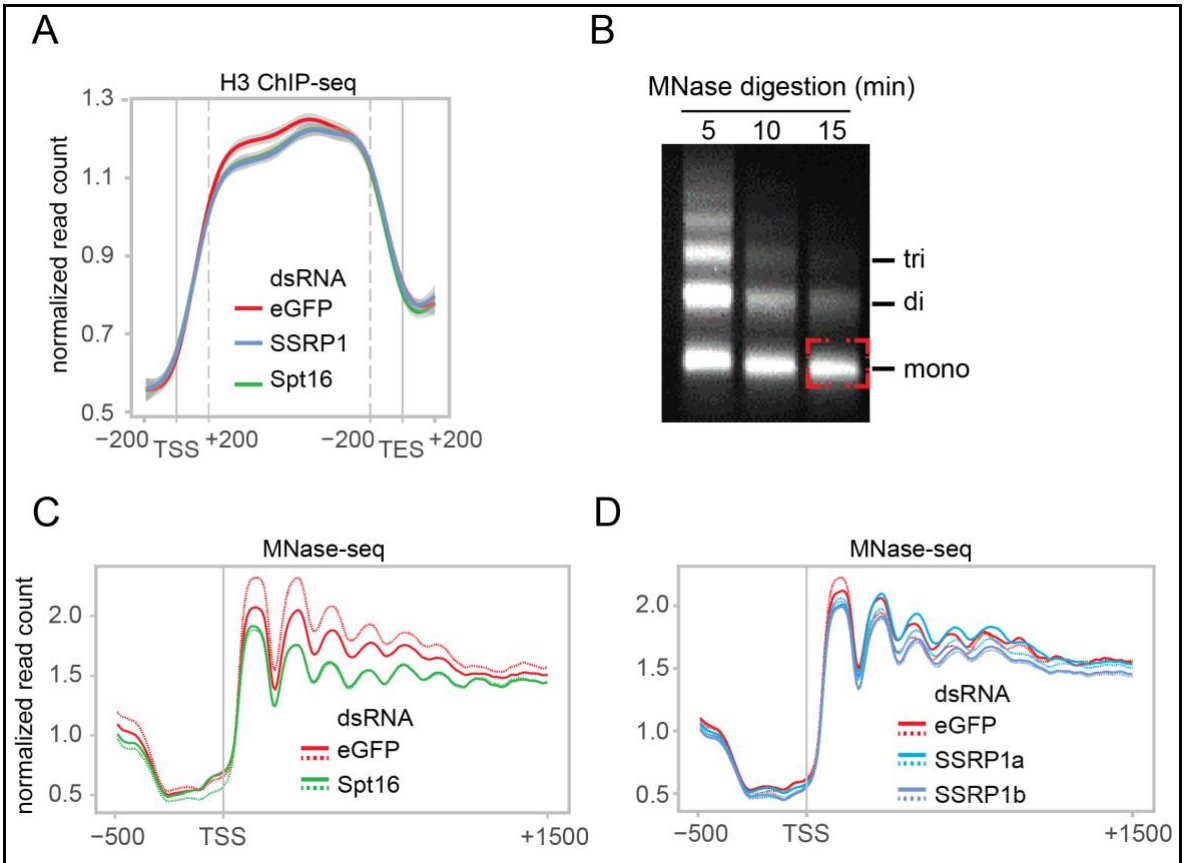

Figure 4. Loss of FACT has minimal effects on nucleosome occupancy and positioning at promoters and transcribed regions globally.
(A) Metagene analysis showing H3 enrichment by ChIP-seq over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), Spt16 dsRNA (marked green), aligned to the TSS and the TES. (B) 1.5% agarose gel showing a time course digestion of chromatin using micrococcal nuclease. Mono-nucleosomal DNA marked in red was excised, purified, and subjected to high-throughput sequencing. (C) Metagene analysis showing nucleosome positioning determined by MNase-seq in cells treated with eGFP dsRNA (marked red) and cells treated with Spt16 dsRNA (marked in green) over transcribed regions aligned to the TSS. (D) Metagene analysis showing nucleosome positioning in cells treated with eGFP dsRNA (marked red) and cells treated with SSRP1 dsRNA (marked in blue) over transcribed regions aligned to the TSS, as determined by MNase-seq. The 2023 non-overlapping genes included in both metagene analyses are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval.

3.3.2 Loss of FACT leads to reduction of promoter H2Av

Since we observed no major change in overall nucleosome positioning and occupancy, we asked whether FACT-dependent changes in gene expression could be

associated with changes in chromatin architecture around promoter regions or elsewhere in and around genes. Recent studies by Jeronimo and colleagues demonstrated that yeast FACT is necessary to restrict a specific variant of H2A called H2A.Z to promoter regions. They discovered that the loss of yeast FACT leads to an increase in H2A.Z within genic regions due to mislocalized SWR-C (Jeronimo et al., 2015). FACT-dependent increase in H2A.Z in genic locations also correlated with an increase in cryptic transcription. This suggests that FACT can regulate transcription outcomes by regulating positioning of H2A.Z. To test this possibility, I performed a ChIP of H2Av (the drosophila orthologue of H2A.Z) in eGFP dsRNA treated and SSRP1 or Spt16 dsRNA treated cells followed by high-throughput sequencing. A metaplot of the distribution of H2Av reads across the transcription start sites and poly-adenylation sites of expressed genes (Figure 5) showed high accumulation of H2Av variant at about 200 bp downstream the TSS which levels off across transcribed regions and towards the poly-adenylation site consistent with published datasets (Bruce et al., 2005; Guillemette et al., 2005; Raisner et al., 2005; Zhang et al., 2005). The average effect of the loss of Spt16 on the position and occupancy of H2Av was only subtle; however, SSRP1-dependent changes in H2A.v at promoters were more pronounced, consistent with evidence that FACT is needed to retain H2A.v at promoter regions in yeast (Jeronimo et al., 2015) .

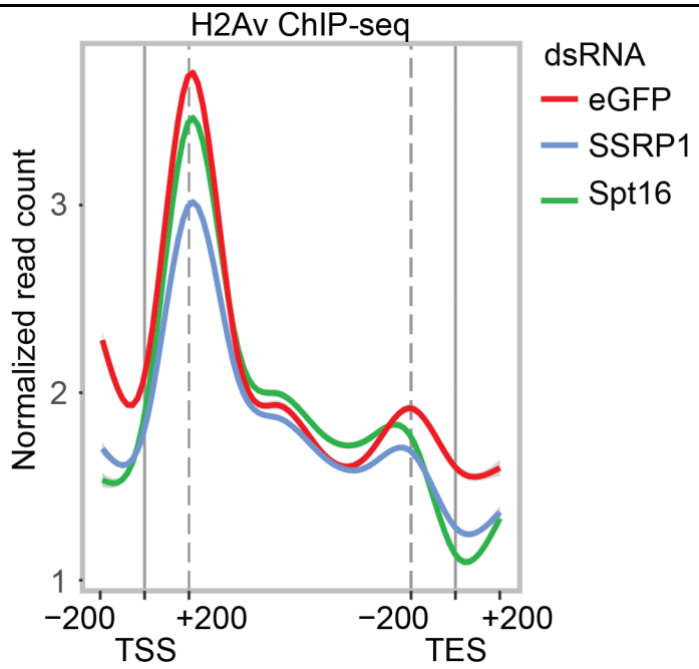

Figure 5. Loss of FACT leads to minimal changes in global H2Av occupancy.
Metagene analysis showing H2Av enrichment by ChIP-seq over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), Spt16 dsRNA (marked green), aligned to the TSS and the TES. The 2023 non-overlapping genes included in both metagene analyses are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval.

3.3.3 Loss of FACT alters patterns of histone modifications associated with promoters of transcribing genes

Over the years, residues on histone tails have been shown to be post-translationally modified by specialized enzymes. These modifications have been shown to be associated with transcription activation, repressed chromatin states, and cell cycle progression. Trimethylation of H3K4 has been shown to be associated with promoters of active genes (Pokholok et al., 2005). H3K4me3 methyl transferase interacts with Ser5P Pol II CTD therefore modifies H3K4 in a transcription dependent manner. Using ChIP followed by qPCR, Stanlie and colleagues demonstrated that FACT can regulate H3K4me3 in Ig class

switch recombination (Stanlie et al., 2010) over Sμ and Sα gene loci. Our observation that loss of FACT leads to deregulation of gene expression led us to ask whether FACT dependent changes in RNA transcript accumulation could be explained by changes in modified chromatin associated with promoters and transcribed genes. To begin to test that, I performed a ChIP of H3K4me3 in eGFP dsRNA treated and FACT depleted cells, followed by high-throughput sequencing. As expected, in control cells, a strong enrichment of H3K4me3 at gene 5' ends was observed in both average gene plots and at individual genes, exemplified in Figure 6B by Inos and 18w. Surprisingly, H3K4me3 signals were drastically reduced at promoters, and increased across transcribed regions in the FACT depleted Drosophila S2 cells. The effect of loss of FACT on H3K4me3 were evident on majority of transcribed genes. To rule out the possibility that the observed FACT-dependent changes in H3K4me3 were not due to FACT-dependent changes in bulk of nucleosomes, we normalized the H3K4me3 reads to H3. The FACT-dependent changes in H3K4me3 were still evident after normalizing to H3 since the FACT-dependent changes in H3 was minimal. Our observations thus far demonstrate that FACT is a key regulator of H3K4me3 and that FACT-dependent changes in H3K4me3 is much more widespread and could possibly explain the changes in transcription. In addition, the 5' -> 3' shift in H3K4me3 following loss of FACT could be consistent with an increase in Pol II release into gene bodies. Since Set1, the H3K4me3 methyltransferase, interacts with Pol II, it is not inconceivable that the 5' -> 3' shift in H3K4me3 is due to FACT-dependent 5' -> 3' shift in Pol II, a hypothesis that is explored in experiments described in Chapter 4.

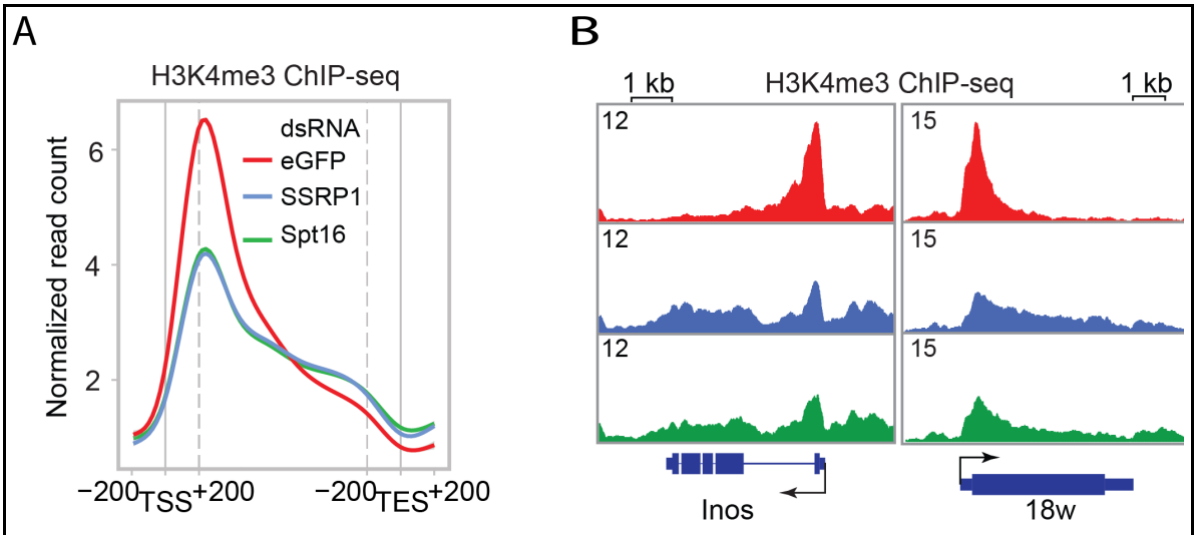

Figure 6. Depletion of FACT alters global distribution of H3K4me3.
(A) Metagene analysis showing H3K4me3 enrichment by ChIP-seq over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), and Spt16 dsRNA (marked in green), aligned to the TSS and the TES. The 2023 non-overlapping genes included in both metagene analyses are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval. (B) IGV genome browser snapshot comparing H3K4me3 signals in cells treated with eGFP dsRNA (blue) to SSRP1 dsRNA (brown) within Inos and 18w loci.

3.3.4 Loss of FACT alters patterns of histone modifications associated with transcribed regions

Next, we asked whether transcriptional changes in FACT knockdown cells because nucleosomes bearing modified histones are not properly reassembled within transcribed regions in the wake Pol II passage. Trimethylation of histone H3 at lysine 36 has been shown to be strongly associated with transcription across gene bodies (Pokholok et al., 2005). In addition, the H3K36me3 methyltransferase, Set2, has been proposed to modulate the recruitment of FACT (Carvalho et al., 2013). I therefore performed a ChIP of H3K36me3 in eGFP dsRNA treated and FACT depleted cells, followed by high-throughput sequencing. Metagene analyses of the read distribution of H3K36me3 across the TSS and the pA show high H3K36me3 occupancy across the gene body with bias towards the

3’end (Figures 7A and 7B) consistent with published datasets (Pokholok et al., 2005). We observed that in the FACT knockdown S2 cells, levels of H3K36me3 were decreased across the gene body compared to the eGFP dsRNA treated cells. Surprisingly, levels of H3K36me3 occupancy were increased beyond the polyadenylation sites of genes (Figures 7A and 7B). These results suggest that FACT is necessary for maintaining proper localization of H3K36me3. Although the mechanism(s) underlying these changes are not known, I speculate that FACT-dependent increases in H3K36me3 beyond the poly-adenylation site could be consistent with failure of Pol II to properly terminate. Furthermore, since Set2 interacts with elongating Pol II, the FACT-dependent 5’ -> 3’ shift in H3K36me3 could be consistent with a FACT-dependent changes in Pol II distribution or increases in transcription elongation rates leading to the shift in histone marks.

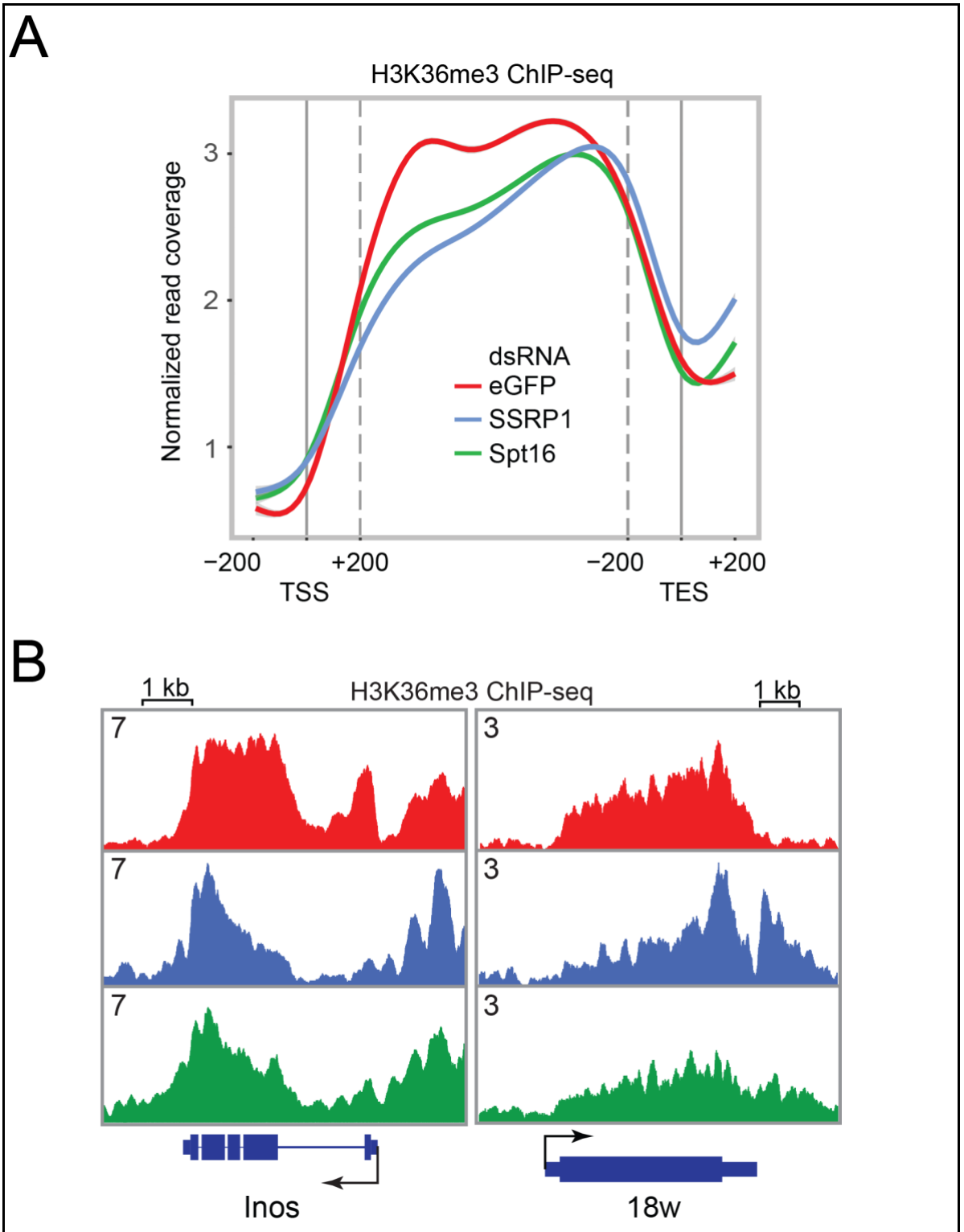

Figure 7. Depletion of FACT alters global distribution of H3K36me3.
Metagene analysis showing H3K36me3 enrichment by ChIP-seq over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), and Spt16 dsRNA (marked in green), aligned to the TSS and the TES. The 2023 non-overlapping genes included in the metagene analysis are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval. (B) IGV genome browser snapshot comparing H3K36me3 signals in cells treated with eGFP dsRNA (colored red) to SSRP1 dsRNA (colored blue) and Spt16 (colored green) within Inos and 18w locus.

3.3.4.1 Newly incorporated histones (H3K56ac) cannot compensate for the FACT-dependent loss of H3K4me3 and H3K36me3

The observations described thus far indicate that FACT plays a major role in maintaining the integrity of modified histones around promoters and transcribed regions during transcription. Upon loss of FACT activity, H3K4me3 and H3K36me3 are shifted in a 5' ->3' direction during transcription. We therefore hypothesized that in order to maintain chromatin stability in FACT knockdown cells, new histones lacking H3K4me3 and H3K36me3 are incorporated by a mechanism independent of FACT activity into nucleosomes in regions where H3K4me3- and H3K36me3-containing nucleosomes have been displaced.

To test this hypothesis, I performed a ChIP of histone 3 lysine 56 acetylation (a mark associated with newly incorporated histones in yeast) followed by high-throughput sequencing in cells that had or had not been depleted of FACT. Metaplots of the read distribution of H3K56ac from 200 bp upstream of TSSs to 200 bp downstream of poly-adenylation sites in both control and FACT-depleted cells showed an accumulation around the promoter with summit at ~200 bp downstream the TSS (Figure 8A and 8B). H3K56ac signals decrease gradually to background levels across the gene body to the pA site. We observed that in the absence of FACT, there was little or no difference in the amount or distribution of H3K56ac after SSRP1 depletion, and only subtle changes in H3K56ac after Spt16 depletion. The apparent slight increase in H3K56ac signal after Spt16 depletion does not appear to be significant, since it is not observed consistently when we plot data from individual replicas of ChIPs performed with chromatin from cells treated

with eGFP dsRNA or the two Spt16 dsRNAs (data not shown). If increases in H3K56 acetylation can indeed be used as a proxy measurement for incorporation of new nucleosomes in *Drosophila*, the finding that there are at most only very modest changes in H3K56 acetylation makes it unlikely that that incorporation of new, unmarked histones could account for the FACT-dependent loss of H3K4me3 around promoters.

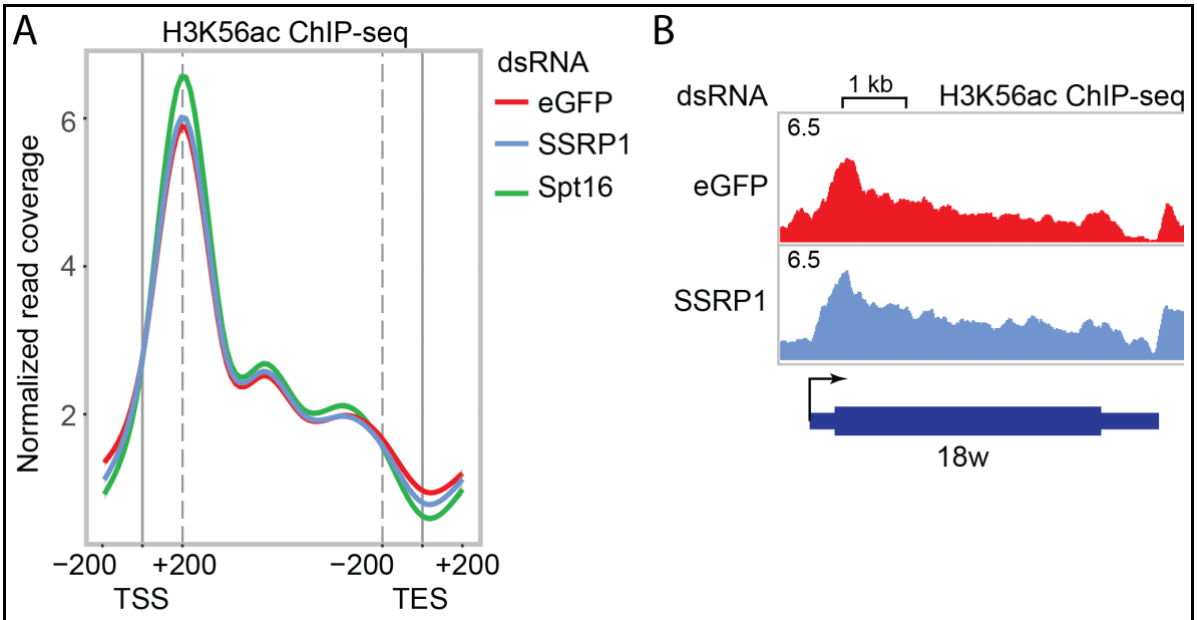

Figure 8. The depletion of FACT has no major changes in global H3K56 ac levels.
(A) Metagene analysis showing H3K56ac enrichment by ChIP-seq over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), and Spt16 dsRNA (marked in green), aligned to the TSS and the TES. The 2023 non-overlapping genes included in both metagene analyses are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval. (B) IGV genome browser snapshot comparing H3K56ac signals in cells treated with eGFP dsRNA (colored red) to SSRP1 dsRNA (colored blue) within 18w locus.

3.4 Discussion

In this chapter, I observed that the depletion of FACT had only minimal effects on histone positioning (measured by MNase-seq) and histone occupancy (measured by H3

ChIP-seq) but alters histone marks associated with transcription. The depletion of FACT subunits resulted in 5’ -> 3’ shifts in H3K4me3 and H3K36me3 but no major change in H3K56 acetylation, which in yeast is a mark on newly incorporated histones. I can envision several models that could explain these findings. First, the 5’ ->3’ shifts in H3K4me3 and H3K36me3 upon depletion of FACT subunits could be because FACT contributes to the proper localization of methyltransferases or demethylases that modify these marks. Alternatively, given that H3K4- and H3K36 methyltransferases interact with transcribing Pol II, it is not inconceivable that the 5’ ->3’ shifts in H3K4me3 and H3K36me3 could be as a result of changes in the distribution of Pol II across genes, perhaps due to differences in overall elongation rate or in the rate of escape of promoter proximally paused RNA polymerase II into more 3' regions of genes.

Chapter 4

Role of FACT in promoter proximal pausing

4.1 Introduction

The genomes of eukaryotes are highly compacted by packaging into nucleosomes to form chromatin. In order for transcription to take place in the cell, RNA polymerase II would have to gain access to promoters, initiate transcription, and elongate nascent transcripts within a chromatin context, where nucleosomes act as barriers limiting the accessibility of DNA sequences in the chromosome. Despite the tight packaging, gene expression is efficient because cells have evolved a collection of chromatin modifying enzymes, ATP dependent chromatin remodelers and histone chaperones to overcome the nucleosome barrier. The histone chaperone FACT is able to stimulate transcription over chromatinized template *in vitro* (Orphanides et al., 1998; Orphanides et al., 1999) by reorganizing nucleosomes. Furthermore, FACT works together with PTEF-b to overcome the basal transcription inhibition imposed by DSIF/NELF *in vitro*. On the other hand, FACT is able to inhibit transcription in a minimal system reconstituted with only TBP, TFIIB, RNAPII and supercoiled DNA template containing the adenovirus major late promoter and a 350 bp G-less cassette (Wada et al., 2000). These studies suggest that FACT can activate or repress transcription *in vitro* through some mechanism. Drosophila polytene chromosome staining of FACT and hyperphosphorylated Pol II showed the two factors co-localize and are redistributed to heat shock genes upon thermal induction (Saunders et al., 2003). Consistent with my evidence that the occupancies of Drosophila FACT and Pol II are correlated, yeast FACT has been reported to travel with Pol II on highly transcribed genes (Mason and Struhl, 2003). In cells deficient of FACT, the levels of the pre-initiation complex (TBP and TFIIB) were reduced at some genes (Mason and Struhl, 2003)

suggesting that not only does FACT regulate transcription elongation but also plays a role in regulating transcription initiation *in vivo*. In higher eukaryotes, the synthesis of mRNA by RNA polymerase is a highly regulated process. Firstly, transcription regulation occurs at the early stages where polymerase together with general transcription factors assemble into the pre-initiation complex at promoters. Secondly, regulation occurs at the early stages of elongation, 30-60 nucleotides downstream of the transcription start site, which is popularly termed promoter proximal pausing. Finally, RNA polymerase regulation occurs after polymerase enters an elongation phase after it escapes pausing checkpoints. Genome-wide studies have shown that promoter proximal pausing of RNA polymerase is a common and one of the key regulatory steps in the transcription of many developmental genes (Adelman and Lis, 2012; Kwak and Lis, 2013). Protein factors like DSIF and NELF have been shown to play a crucial role in maintaining RNA polymerase II at paused regions. Both proteins were shown *in vitro* to inhibit elongation and to co-occupy at paused regions (Wu et al., 2003; Yamaguchi et al., 2013). It was further shown that, depletion of these factors alleviates pausing *in vivo*. Promoter proximal pausing decisions can also be influenced by the core promoter elements (Kwak et al., 2013; Li and Gilmour, 2013) and the +1 nucleosome (Kwak et al., 2013; Li and Gilmour, 2013; Weber et al., 2014). In order for RNA Pol II to escape the promoter proximal region, the kinase P-TEFb comprising of Cyclin T1 and CDK9 phosphorylates DSIF, NELF and the CTD Ser2 of the Pol II subunit Rpb1 to relieve the Pol II of the inhibition (Peterlin and Price, 2006). Once RNA Pol II reaches elongation phase, the rate of elongation can be affected by different, not necessarily mutually exclusive factors. Some DNA sequences may be more difficult to transcribe than others owing to their DNA topology. Histone marks can tighten or loosen DNA binding around nucleosomes, restricting or facilitating Pol II transcription efficiency. Elongation factors, histone chaperones and nucleosome remodelers can facilitate RNA Pol

II movement through chromatin by alleviating pausing and stalling during elongation, either by altering properties of elongating Pol II or of chromatin architecture (Jonkers and Lis, 2015).

The role of FACT in regulating transcription initiation and elongation is not well understood in higher eukaryotes where transcription programs are complex. In this chapter, we provide evidence that is consistent with the role of FACT in promoter proximal pausing.

4.2 Materials and Methods

4.2.1 Chromatin Immunoprecipitation of Rpb3

4.2.1.1 Cell culture, dsRNA mediated knockdown of FACT and western blot

ChIP of Rpb3 was performed on Drosophila S2 cells depleted of FACT subunits as described above.

4.2.1.2 Formaldehyde crosslinking, sonication and immunoprecipitation

Formaldehyde crosslinking and sonication were carried out as described above using 10 µg of anti-Rpb3 (W. Shao and J. Zeitlinger) antibody coupled to Dynabeads® protein G (life technologies, 10004D).

4.2.1.3 DNA purification

The reverse crosslinked ChIP DNA was purified using Phenol:Chloroform:Isoamyl Alcohol as described above.

4.2.1.4 Library preparation and DNA sequencing

The Molecular Biology Core of the Stowers Institute prepared libraries from the ChIPed DNA using the KAPA HTP kit (Kapa Biosystems), sized selected using the SPRIselect double side size selection kit (Beckman Coulter) and performed DNA sequencing on the Illumina HiSeq platform.

4.2.1.5 ChIP-seq analysis

The ~30million reads from each experiment was aligned using bowtie 2 to the drosophila dm6 genome as described above.

4.2.2 PRO-seq (precision nuclear run-on and nuclear)

4.2.2.1 Cell culture, dsRNA mediated knockdown of FACT and western blot

In order to study the role of the FACT complex on transcriptionally engaged polymerases, I performed a PRO-seq on Drosophila S2 cells depleted of FACT subunit, SSRP1. For each replica, about 12×10^6 cells were seeded in a 100 mm dish and incubated at 25°C for 30 min to allow cells to attach to the bottom of the plate. Cells were then treated by adding 60 µg dsRNA to each dish and incubated for 48 hours at 25 degrees Celsius. A second dose of 60 µg dsRNA/dish was added and the cells harvested after 48 hours to study the loss of function phenotypes. Cells were harvested and washed once in 10 ml 1X PBS by centrifugation at 1000 x g for 4min in a swinging bucket at 4°C. A western blot was performed on the non-targeting control and FACT depleted cells to confirm SSRP1 knockdown.

4.2.2.2 Nuclear isolation for PRO-seq

PRO-seq procedure and library preparation were performed exactly as previously described (Kwak and Lis, 2013; Mahat et al., 2016). Briefly, nuclei were isolated from the washed cells by resuspending gently in 5ml ice cold douncing buffer (10 mM Tris-HCl pH 7.4, 300 mM sucrose, 3 mM CaCl2, 2 mM MgCl2, 0.1% Triton X-100, 0.5 mM DTT, 1 tablet of protease inhibitors cocktail (Roche 11873580001) per 50ml, 4 u/ml RNase inhibitor (SUPERaseIN, Ambion AM2696), incubated on ice for 5 min and dounced until 90% of cells were lysed (from 5-25) in a 5ml capacity Dounce homogenizer. Nuclei were then pelleted at 1000 x g for 4 min in a swinging bucket at 4°C, washed twice using 10 ml douncing buffer and resuspended in storage buffer (10 mM Tris-HCl pH 8.0, 25% (vol/vol) glycerol, 5 mM MgAcetate 0.1 mM EDTA, 5 mM DTT) at a concentration of $20x10^6$ nuclei/100 µl. The isolated nuclei in storage buffer were flash frozen in liquid nitrogen and stored at -80°C.

4.2.2.3 4-biotin nuclear run-on and PRO-seq library preparation

To perform the 4-biotin run-on assay, $20x10^6$ nuclei in 100 µl storage buffer were mixed thoroughly with an equal volume of preheated (30°C) 2X Nuclear Run-On master mix (10 mM Tris HCl pH 8.0, 5 mM $MgCl_2$, 300 mM KCl, 1 mM DTT, 0.8 u/µl RNAse inhibitor (Ambion, AM2696), 1 % Sarkosyl and 50 µM each of all 4 biotin-NTPs (PerkinElmer Biotin-11-ATP, NEL544001EA; Biotin-11-GTP, NEL545001EA; Biotin-11-CTP NEL542001EA: Biotin-11-UTP NEL543001EA) and incubated for 3 min at 30°C in a heat block. The nascent RNA was extracted using Trizol LS (Invitrogen, 10296-028) and precipitated using 2.5 volumes of 100% ethanol. The purified RNA was resuspended in 20 µl DEPC water, heat denatured at 65°C for 40 s and placed on ice. Denatured RNA samples were then fragmented by adding 5 µl of 1N NaOH and incubated on ice for 10

min, and neutralized by adding an equal volume of 1 M Tris-HCl pH 6.8. Excess salt and residual NTPs were then removed by passing base-hydrolyzed RNA samples through the P-30 column (Bio-Rad, 732-6250). In order to protect RNA samples from degradation, 1 µl RNase inhibitor (Life Technologies, AM2696) was added to each sample. Prior to enriching RNAs for biotinylated transcripts, 90 µl of streptavidin-coated magnetic beads (Life Technologies, 11206D) per library were washed once in 100 µl of buffer containing 0.1 N NaOH and 50 mM NaCl in DEPC water, and twice using 100 mM NaCl in DEPC water. Beads were finally resuspended in 150 µl binding buffer and divided into 3 aliquots. Biotinylated RNA transcripts were then enriched by binding fragmented RNAs to pre-washed streptavidin M280 beads (Invitrogen, 11206D). Briefly, about 50 µl of RNA samples were bound to 50 µl pre-washed Streptavidin beads and incubated for 20 min at room temperature on a rotator. Bead-bound RNA samples were then washed twice in 500 µl ice-cold high salt wash buffer (2 M NaCl, 50 mM Tris-HCl pH 7.4, 0.5% Triton X-100), twice in 500 µl binding buffer (300 mM NaCl, 10 mM Tris-HCl pH 7.4, 0.1 % Triton X-100), and once in 500 µl low salt buffer (5 mM Tris-HCl pH 7.4, 0.1% Triton X-100). Finally, RNAs were extracted from the beads using Trizol (Invitrogen, 15596-018) and ethanol precipitation. In order to ligate the 3’ RNA adaptor (5Phos/rGrArU rCrGrU rCrGrG rArCrU rGrUrA rGrArA rCrUrC rUrGrA rArC/3Inverted dT; IDT custom synthesis, RNase-free HPLC purified) to the fragmented RNAs, RNAs were resuspended in 4 µl 3’ RNA adaptor dilution (0.5 µl of 100 µM 3’ RNA adaptor in 3.5 µl of DEPC water), heat denatured at 65 °C for 20 s and placed on ice. Denatured RNA samples were mixed with 6 µl RNA ligation mix (1 X T4 RNA ligase buffer, 1 mM ATP, 10 % PEG, 4 u/µl RNase inhibitor and 1 u/µl T4 RNA ligase I NEB, M0204S) and incubated at 20 °C for 4 h. To remove excess adaptors and salts from the ligated RNA, second biotin enrichments were performed using streptavidin beads as described above. RNAs were then purified from streptavidin beads using Trizol

and ethanol precipitation. Prior to ligating the reverse 5' RNA adaptor (rCrCrU rUrGrG rCrArC rCrCrG rArGrA rArUrU rCrCrA, Custom synthesis from IDT, RNase-free HPLC purified) to the RNA, RNA pellets were resuspended in 5 µl DEPC water, heat denatured briefly at 65 °C for 20 s and placed on ice. Heat-denatured short nascent RNAs were then de-capped by mixing with 5 µl 5'cap repair enzyme mix (1X ThermoPol Reaction Buffer, 2 u/µl RNase inhibitor, 0.5 u/µl RNA 5' pyrophosphohydrolase, RppH, NEB, M0356S) and incubating at 37°C for 1 h. Next, the decapped RNAs were mixed with 90 µl PNK mix (1 X PNK buffer, 1 mM ATP, 1 u/µl RNase inhibitor and 0.25 u/µl T4 Polynucleotide Kinase NEB, M0201S) and incubated at 37°C for 1h. The 5'end modified RNAs were then purified using Trizol LS (Invitrogen, 10296-028) followed by ethanol precipitation. To perform the 5' adaptor ligation, RNAs were dissolved in 4 µl of 5' RNA adaptor dilution (0.5 µl of 100 µM 3' RNA adaptor in 3.5 µl of DEPC water), heat-denatured at 65°C for 20 s and placed on ice. RNAs were then mixed with 6 µl RNA ligation mix (1 X T4 RNA ligase buffer, 1 mM ATP, 10 % PEG, 4 u/µl RNase inhibitor and 1 u/µl T4 RNA ligase I NEB, M0204S) and incubated at 20°C for 4 h. The cloned RNA products were then subjected to another round of biotin enrichment and purification using Trizol followed by ethanol precipitation and resuspended in 10 µl of DEPC water. To reverse transcribe the cloned products, RNAs were mixed with 2.5 µl RT primer mix (2.5 µM RP1 primer (AAT GAT ACG GCG ACC ACC GAG ATC TAC ACG TTC AGA GTT CTA CAG TCC GA, Custom synthesis from IDT, RNase-free HPLC purified), 625 µM dNTP mix) and incubated at 72°C for 2 min and chilled on ice for 2 min. 6 µl of RT buffer mix (1 X First strand buffer, 5 mM DTT, 2 u/µl RNase inhibitor) was added to each RNA-primer mix and incubated at 37°C for 5 min. 1.5 µl Superscript III reverse transcriptase (Invitrogen, 56575) were mixed with the RNAs and incubated for 15 min at 45 °C, then 40 min at 50 °C, 10 min at 55 °C and 15 min at 70 °C. Finally, 6 µl of DEPC water was added to the RT reaction.

An aliquot of the 26 µl cDNA from each replicate was used to carry out trial amplifications to determine the optimal number of cycles in order to avoid over amplification of the library. Four-fold serial dilutions (2 µl cDNA + 6 µl water) of cDNA were prepared using 2 µl cDNA aliquot from each experiment. Trial PCRs were carried out by mixing 1 µl of each serially diluted cDNA samples with 14 µl PCR mix (1 X HF buffer, 1 M betaine, 250 µM each dNTP mix, 250 nM RP1 primer, 250 nM of either RPI4, RPI5, RPI6, or RPI7, and 0.02 u/µl Phusion DNA polymerase, NEB, M0530S) and running samples on the Eppendorf Mastercycler epgradient S using the following thermal conditions: 95 °C for 2 min, 5 cycles of 95 °C, 56°C, 72 °C each for 30 s, 18 cycles of 95 °C, 65 °C, 72 °C each for 30 s, followed by 72°C for 10 min, and stored at 4°C. Primers RPI4 (RNA PCR Primer Index 4, CAA GCA GAA GAC GGC ATA CGA GAT TGG TCA GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA, Custom synthesis from IDT, RNase-free HPLC purified), RPI5 (RNA PCR Primer Index 5, CAA GCA GAA GAC GGC ATA CGA GAT CAC TGT GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA, Custom synthesis from IDT, RNase-free HPLC purified), RPI6 (RNA PCR Primer Index 6 CAA GCA GAA GAC GGC ATA CGA GAT ATT GGC GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA, Custom synthesis from IDT, RNase-free HPLC purified) and RPI7 (RNA PCR Primer Index 7, CAA GCA GAA GAC GGC ATA CGA GAT GAT CTG GTG ACT GGA GTT CCT TGG CAC CCG AGA ATT CCA, Custom synthesis from IDT, RNase-free HPLC purified) were used to barcode cDNAs from NS1 (Non-targeting control, replicate 1), NS2 (Non-targeting control, replicate 2), SSRP1kda1 (SSRP1 dsRNA mediated knockdown, replicate 1), and SSRP1kda2 (SSRP1 dsRNA mediated knockdown, replicate 2) respectively. PCR amplicons from trial amplification were run on the Bioanalyser to determine the dilution with the optimum conditions (sufficient amount of product, not over amplified and having 50-75% of unused primers). Final full-scale PCR amplifications were carried out by mixing the remaining 24 µl cDNA samples with 0.5 µl RPI-n (25 µM)

primer, 25.5 µl full scale primer mix (1 X HF buffer, 1 M Betaine, 250 µM each dNTP mix, 250 nM RP1 primer, 0.04 u/µl Phusion DNA polymerase), and running samples on the Eppendorf Mastercycler epgradient S using the following thermal conditions: 95 °C for 2 min, 5 cycles of 95 °C, 56°C, 72 °C each for 30 s, 13 cycles of 95 °C, 65 °C, 72 °C each for 30 s, followed by 72°C for 10 min, and stored at 4°C. PCR products were purified using ethanol precipitation and resuspended in 18 µl of water. PCR amplicons running from 140 bp to 350 bp were selected using the 2% Agarose Gel Cassette, Dye-Free (Sage Science, CDF2010, Cassette type: 2% DF Marker L) on the Pippen Prep™ (Sage Science, Software: v.5.8) instrument. Briefly, DNA samples were made up to 30 µl with 1 X TE (10 mM Tris, 1 mM EDTA, pH 8.0) buffer and mixed thoroughly with 10 µl of loading solution/marker. Next, the Pippin Prep software was launched, and the 2% DF Marker L protocol selected. Internal standards options were selected, and size selection parameters were set to 140 bp – 350 bp. Prior to running, the Pippen optics were calibrated by placing the calibration fixture onto the optical nest and pressing calibrate to launch the calibration window. 0.80 was entered in the “Target I ph, mA” and the calibrate button was pressed. After the calibration was complete the exit button was pressed. Next, buffer levels in the 2% gel cassette wells were inspected. Gel cassette was placed into the optical nest and the white adhesive strip removed. Buffers in the elution modules were removed and replaced with 40 µl of fresh electrophoresis buffer. Elution wells were then sealed with an adhesive tape strip. Prior to loading the library, sample wells were filled to the top with buffer. A continuity test was performed to measure the current in each separation and elution channel. If everything goes well, a “PASS” should be returned for each separation and elution channel. To load DNA samples, 40 µl of buffer were removed from each well and 40 µl of the prepared sample DNA loaded. To run, the “START” button was pressed on the main protocol tab. At the end of the run, samples were collected into 1.5 ml tubes and

quantified using the bioanalyzer. Equimolar concentrations of library fractions were pooled together and sequenced using a mid-output flow cell on the Illlumina NextSeq platform.

4.2.2.4 PRO-seq bioinformatic analysis

In all cases, we used the Dmel Release dm6 gene and transcript annotations downloaded from ensemble. The adapter sequences 5'-TGGAATTCTCGGGTGCCAAGG-3' were removed from the reads and the reads trimmed to a maximum length of 36 bp. All reads less than 15 bp were filtered out, reverse-complemented and aligned with Bowtie2 to dm6. RPKM were calculated as counts in each window per total mapped reads per length of the window.

4.2.2.5 Active transcript identification

In order to provide an overview of various genomic signatures for a given set of genes, we algorithmically assigned a specific "most active" transcript isoform to all known genes. To do this we used PRO-Seq promoter coverage, defined as -100 to +300 bp around all annotated transcription start sites (TSS), as a proxy for active transcripts. For each gene, the TSS from the site with the highest promoter coverage was kept for downstream analysis.

4.2.2.6 Promoter pausing ratio (PPR)

Pausing ratio was calculated as described (Kwak et al., 2013). Any genes that were less than 500 bp in length, overlapped other genes or were within 200 nt of nearby genes were excluded from the downstream analysis. The Pol II pause release ratio PRR was calculated by dividing the average gene body coverage (defined as +300 from the TSS to the end of the transcript) by the average promoter coverage (defined as -50 to +150

around the TSS). We also excluded genes with less than 1 RPKM PRO-seq signal in the promoter region in order to exclude non-active genes. In making this assumption we understand that in some cases greater coverage could be also be related to Pol II pausing and that alternative start sites within 100 nt of each other may result in confounding TSS selection.

4.3 Results

4.3.1 Loss of FACT leads to a decrease of Pol II at promoters of most genes and a subsequent increase gene bodies of a subset of genes

Observations made thus far raise the possibility that FACT could play a role in regulating transcription initiation, elongation, or both. To investigate the role of FACT in regulating transcription initiation and elongation, I began by performing a ChIP-seq of Rpb3 in cells depleted of FACT and eGFP dsRNA treated cells to map RNA polymerase II independent of the phosphorylation state of the CTD. In addition, base-pair resolution of transcriptionally engaged polymerases were mapped, genome-wide, using the newly developed nuclear run-on assay called PRO-seq in *Drosophila* S2 cells depleted or not of FACT. An average gene plot of the distribution of ChIP signals of total RNA polymerase II (Rpb3 subunit) across the transcription start site (TSS) and the poly-adenylation site (pA) showed, as expected, an accumulation of Pol II at the TSS and a signal just higher than background across the transcribed regions in the non-targeting control S2 cells (Figure 9A). We also observed that average gene plots of transcriptionally engaged RNA polymerase II (Figure 9B) measured using PRO-seq had very similar profile as the Pol II ChIP but with a base pair resolution. As shown in Figures 9A, 9B, 9C, and 9D RNAi-mediated loss of FACT led to a decrease in Pol II accumulation at the TSS and an increase in RNA polymerase II signal across gene bodies of many non-overlapping genes (the 2023

genes included in the metagene analyses are longer than 0.5 kb, and have no overlapping genes or flanking genes within 200 bp to minimize ChIP signals associated with nearby genes). This observation suggests that FACT plays a role in regulating RNA polymerase II accumulation across transcribed regions. Our data are also consistent with the models that FACT is necessary for the proper loading of Pol II and/or helps to slow the transition of Pol II from promoter-proximal pausing to productive elongation.

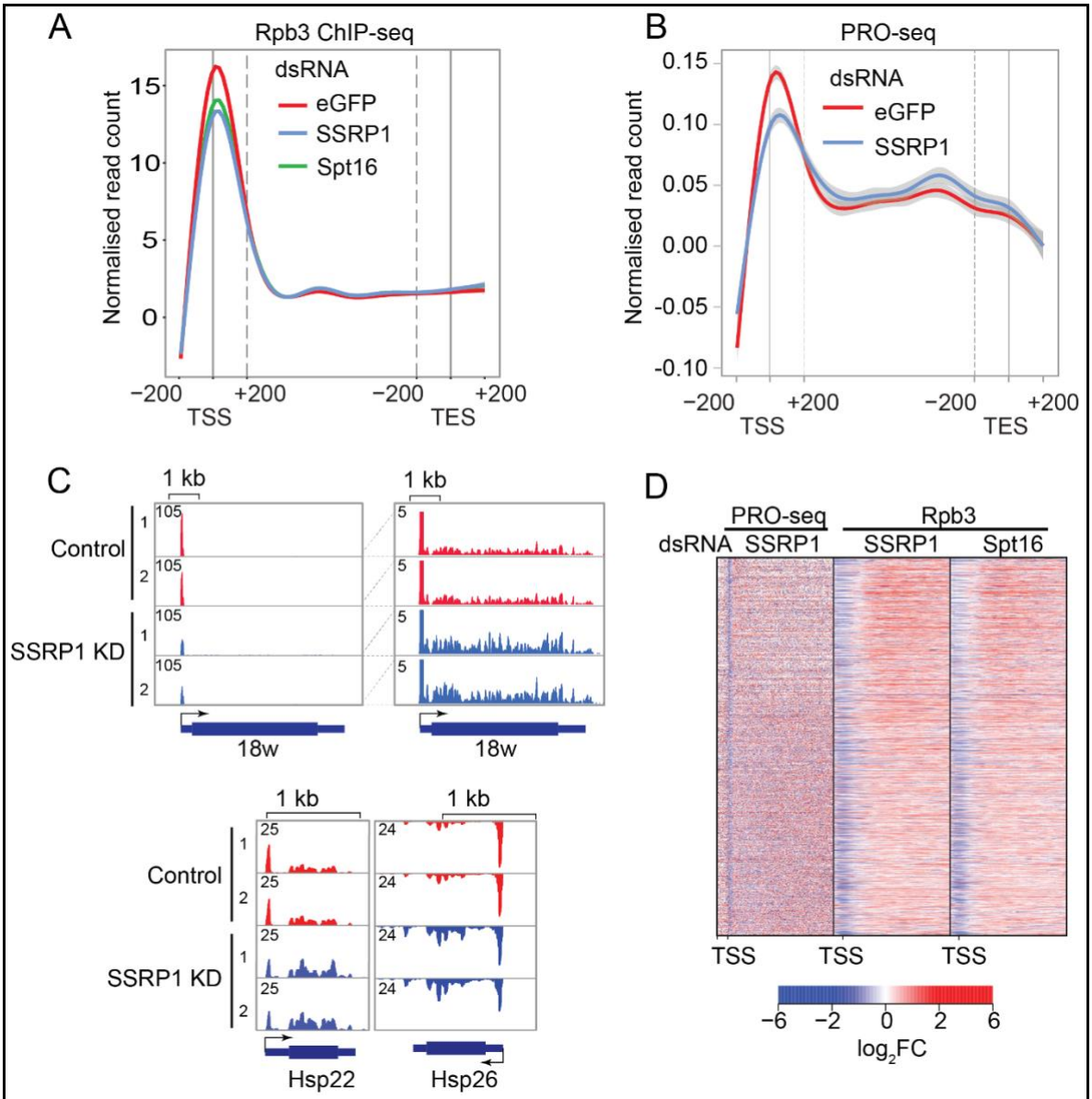

Figure 9. FACT depletion alters distribution of transcriptionally engaged polymerases.
(A) Metagene analysis showing Rpb3 enrichment by ChIP-seq over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), and Spt16 dsRNA (marked in green), aligned to the TSS and the TES. (B) Average PRO-seq profile over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue) aligned to the TSS and the TES. (C) IGV genome browser snapshot comparing PRO-seq signals in cells treated with eGFP dsRNA (colored red) to cells treated with SSRP1 dsRNA (colored blue) within 18w (a highly paused gene), Hsp22 (locus highly bound by FACT), and Hsp26 (locus highly bound by FACT) loci. Top right panel showing a rescaled version to emphasize changes in signal in the gene body. (D) Heatmap showing fold changes in PRO-seq and Rpb3 signals in cells depleted of either subunit of FACT as indicated above. Brightest blue represents a fold decrease of 64 while the brightest red a fold increase of 64
The 2023 non-overlapping genes included in the metagene analysis are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval.

4.3.2 Potential role of FACT in promoter proximal pausing

The 5' accumulation of RNA polymerase II genome-wide in eukaryotes has previously been reported as evidence for promoter proximal pausing (Gariglio et al., 1981; Rougvie and Lis, 1990; Wade and Struhl, 2008). The widely held view in the field is that promoter proximal pausing is a key rate limiting step in transcription elongation that can be caused by cis-regulatory elements and transacting factors. Thus, our observation that loss of FACT results in the decrease of promoter proximal Pol II signal led us to ask whether FACT plays a role in pausing. To address that, we calculated pausing indices using data from PRO-seq, which has some advantages over the ChIP-seq of Pol II. Advantages of using PRO-seq over ChIP-seq of Pol II are that (i) PRO-seq provides base-pair resolution, (ii) it provides strandedness information, and (iii) it measures only transcriptionally engaged polymerases. We then calculated the pausing index (Kwak et al., 2013), which is the ratio of Pol II signal in the promoter proximal region (from 50 bp upstream the TSS to 150 bp downstream the TSS) versus the gene body (from 300 bp downstream the TSS to the polyadenylation site) (Figure 10A). We then classified the non-overlapping genes (N=2023) into highly paused (pausing index ≥ 23.5; n = 674 genes) moderately paused (3.3 < pausing index < 23.5; n = 674 genes) and lowly paused (pausing index ≤ 3.3; n = 674) based on their pausing indices (Figures 10A and 10B). A metaplot of the distribution of total and transcriptionally engaged polymerases showed a high, medium and low promoter proximal signal in the highly, moderately and lowly paused categories respectively (Figure 10B). We then calculated the pause release ratio (Chen et al., 2015a) which is the inverse of our pausing index. Next, we calculated the paused release ratio fold change (PRR_FC) after FACT knockdown to explore the effect of loss of FACT on highly, moderately and lowly paused genes. A boxplot analysis of the paused release ratio fold change in the three pausing categories show greater increases in pause-release ratio

in more highly paused genes (Figure 10C). This data suggest that the loss of FACT has strongest effects on genes that are highly paused genes. In addition, the increase in promoter proximal release fold change among highly paused genes following loss of FACT is consistent with an increase in the rate of escape of promoter proximal Pol II and increase in distribution of Pol II in gene bodies

4.3.3 Loss of FACT leads to increased RNA transcript abundance of highly paused genes

Release from promoter proximal pausing is believed to be a rate limiting step in the transcription of many genes. If the observed increase in PRR at paused genes indeed reflects an increase in the rate with which Pol II is released into productive elongation, we might expect that FACT knockdown would lead to an increase in the abundance of mature transcripts at genes that exhibit the greatest changes in PRR. To investigate this possibility, we performed a boxplot analysis of RNA transcript abundance fold change in the 3 pausing groups defined above (Figure 10D). Consistent with previous findings, t highly paused genes have relatively lower transcript abundance compared to moderately and lowly paused genes. in addition, following SSRP1 knockdown, we observed a significant increase ($p< 1.78 \times 10^{-11}$) in steady state RNA transcript abundance among the highly paused genes compared to lowly paused genes.

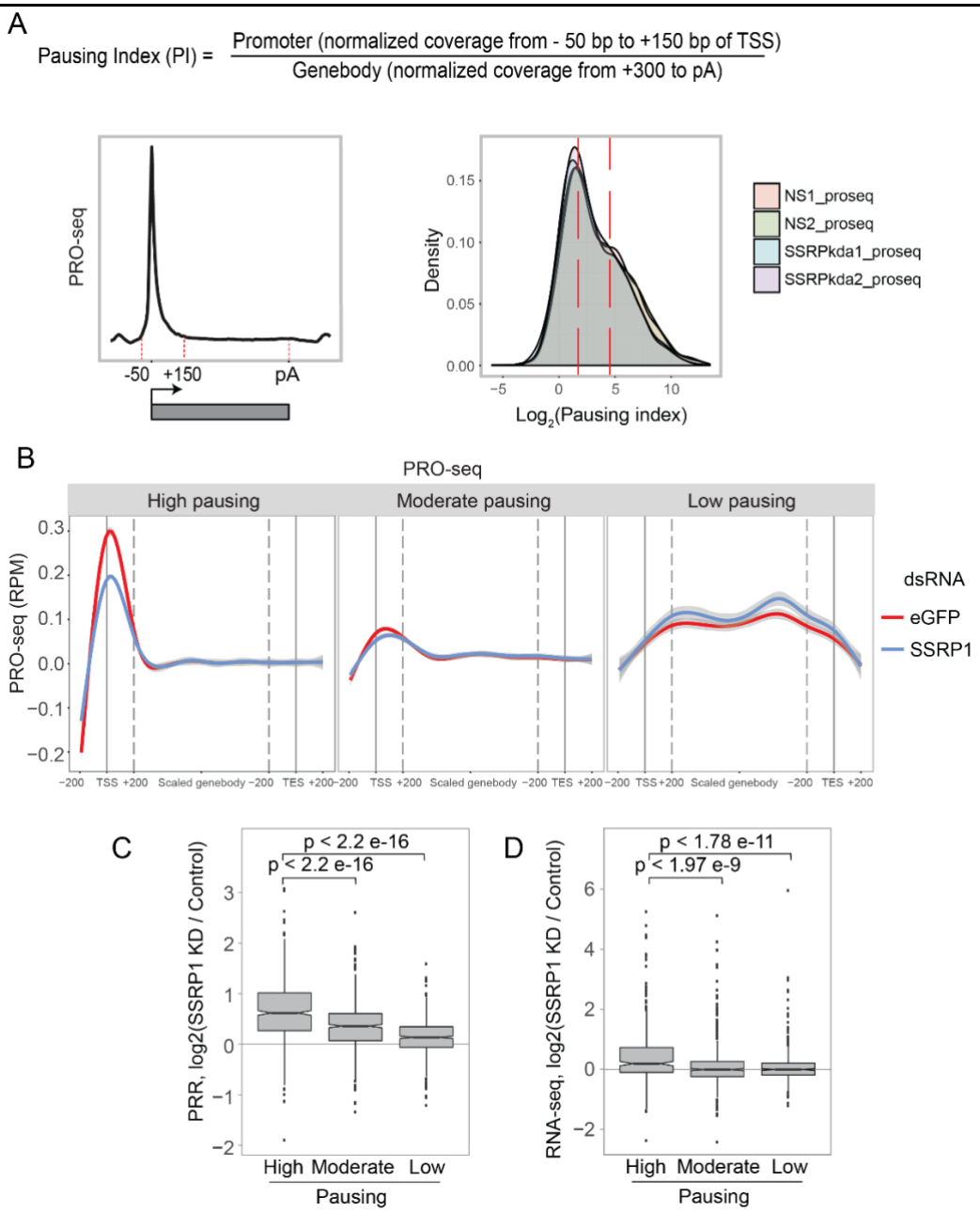

Figure 10. Highly paused genes are the most sensitive to the loss of FACT.
(A) Schematic representation of the pausing index formula. Panel on the right showing distribution of pausing indices across eGFP dsRNA treated (indicated as NS1 and NS2) and SSRP1 (indicated as SSRP1kda1 and SSRP1kda2) deficient S2 cells. (B) Metagene analyses showing transcriptionally engaged RNA polymerase II (PRO-seq) across genes with varying degrees of pausing in eGFP dsRNA treated (marked red) and SSRP1 dsRNA treated (marked in blue) cells, aligned to the TSS and the TES. The non-overlapping genes included metagene analyses are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval. (C) Boxplot analysis of pause release ratio log2 fold change following SSRP1 knockdown across genes with varying degrees of pausing. (D) Boxplots analysis showing log2 fold change in RNA transcript abundance following SSRP1 knockdown across genes with varying degrees of pausing.

4.3.4 Knockdown of FACT has stronger effects on H2Av, H3K4me3 and H3K36me3 on low paused genes than highly paused genes

Based on observations thus far, I hypothesized that the degree to which the modified histones or the chromatin architecture can be reorganized (disassembled and /or reassembled) by FACT would be dependent on the level of pausing of the gene. To test this hypothesis, we performed metagene analyses of H3, MNase, H2Av, H3K4me3, H3K36me3 and H3K56ac across genes with varying degrees of pausing (highly paused genes, moderately paused genes and lowly paused genes) in eGFP dsRNA treated and FACT depleted cells (Figure 11-12).

Results of this analysis indicate that lowly paused genes have higher nucleosome occupancy than highly paused genes (Figure 11A), consistent with observations made by the Adelman lab (Gilchrist et al., 2010). The loss of either subunit of FACT had no major change on H3 or H3K56ac occupancy on either highly paused or lowly paused genes (Figure 12C). Also, consistent with previous observation, nucleosomes on lowly paused genes tend to be in an ordered state compared to highly paused genes. Again, the loss of either subunit of FACT has no major change in MNase-seq profiles on either highly paused or lowly paused genes (Figure 11B).

Occupancy of the histone variant H2Av and of the modified histones H3K4me3 and H3K36me3 was greatest on genes that exhibit low pausing and highest on the high pausing genes. The loss of FACT, and in particular of its SSRP1 subunit, led to more substantial decreases in H2Av occupancy at low and moderately paused genes, but had minimal effect on H2Av occupancy at highly paused genes (Figure 11C). In addition, FACT dependent loss or 5’ ->3’shift in H3K4me3 and H3K36me3 were greatest among genes that are lowly paused compared to highly and moderately paused genes Figure 12A and 12B).

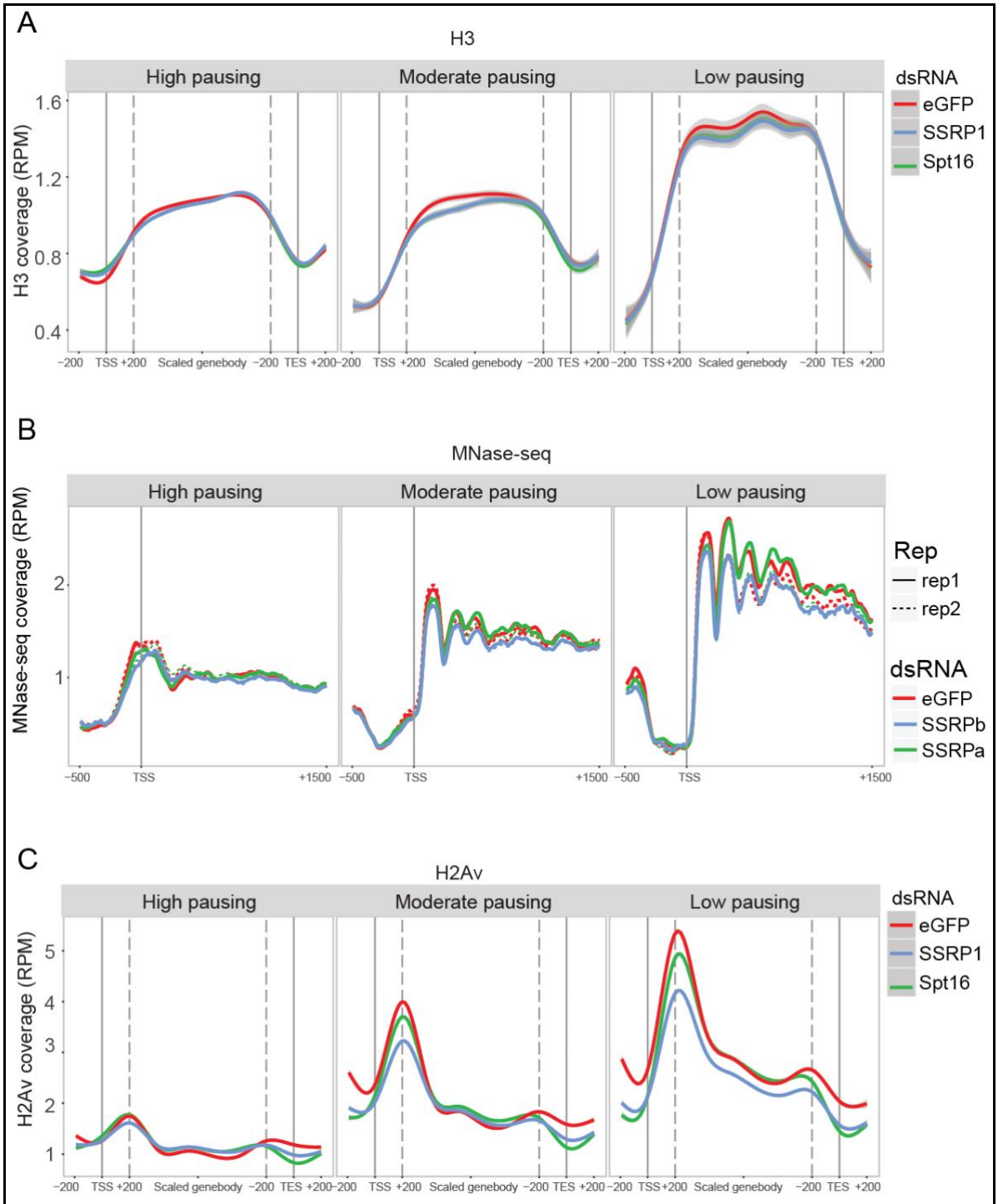

Figure 11. The loss of FACT has minimal effects on nucleosome occupancy, nucleosome positioning and H2Av independent of the pausing status of genes.
Metagene analysis showing (A) H3 ChIP, (B) Nucleosome positioning measured by MNase-seq, (C) H2Av over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), Spt16 dsRNA (marked green), aligned to the TSS and the TES. The 2023 non-overlapping genes included in each metagene analyses are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval.

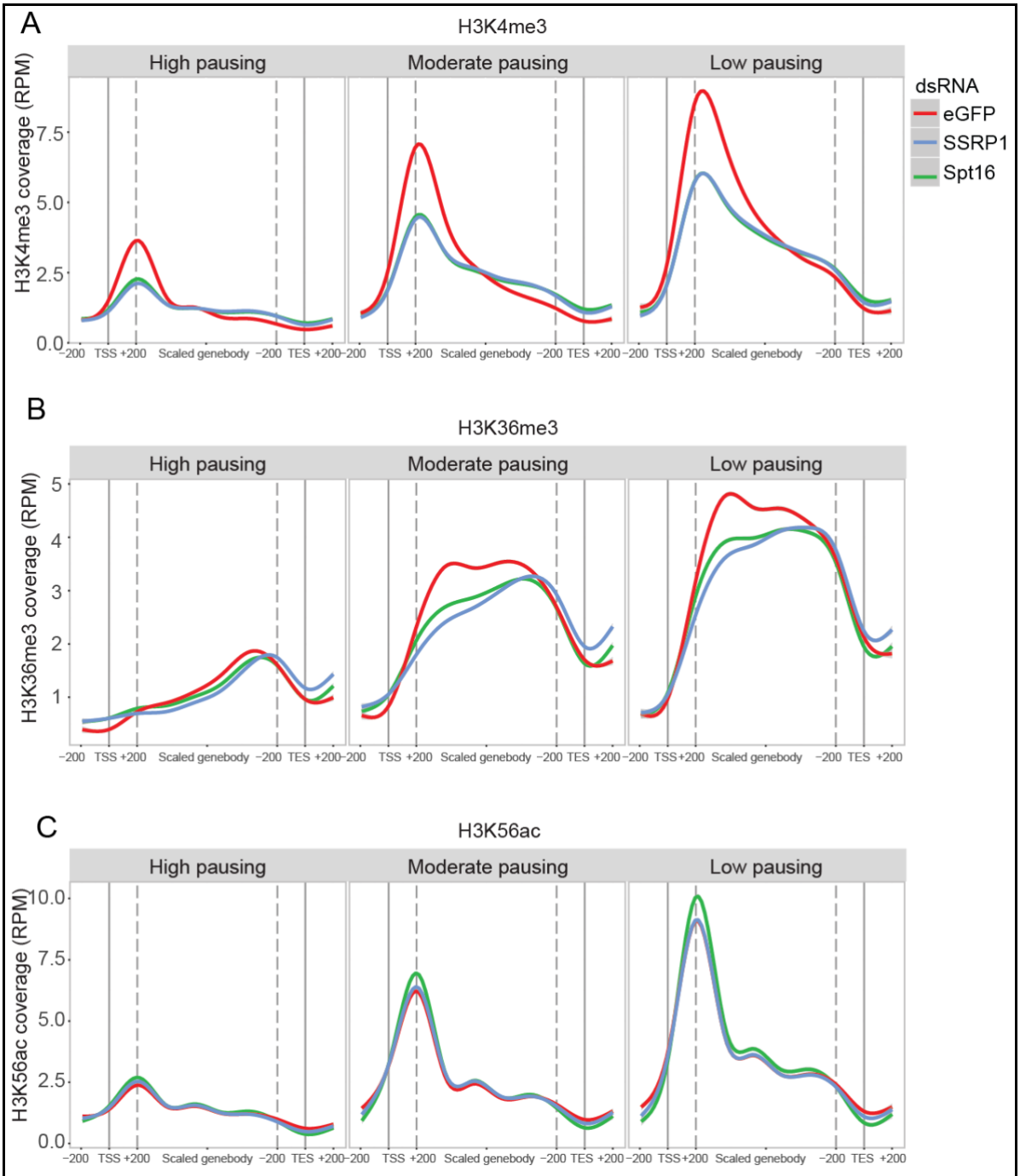

Figure 12. FACT depletion affects H3K4me3 and H3K36me3 with minimal effects on H3K56ac across genes with varying degrees of pausing.
Metagene analysis showing (A) H3K4me3, (B) H3K36me3 (C) H2Av enrichment by ChIP-seq over transcribed regions in Drosophila cells treated with eGFP dsRNA (marked in red), SSRP1 dsRNA (marked in blue), Spt16 dsRNA (marked green), aligned to the TSS and the TES. The 2023 non-overlapping genes included in each metagene analyses are active (have normalized promoter PRO-seq signal greater than 1 RPKM), longer than 0.5 kb and have no neighboring genes within 200 bp to reduce signals associated with nearby genes. Shaded area represents 95% confidence interval.

4.4 Discussion

The contribution of FACT in regulating the interplay between transcription and histone modification is not well understood in higher eukaryotes where transcription programs are complex. In this chapter, I provide data consistent with a role of FACT in promoter proximal pausing.

Using high-throughput methods, we observed that the depletion of FACT leads to significant reduction of transcriptionally engaged RNA polymerase II around promoters, accompanied by an increase in transcriptionally engaged Pol II in gene bodies of a number of genes. Of note, highly paused genes exhibited the greatest dependence on FACT since we observed the greatest change in pause release ratio fold change among this class of genes. The observed FACT dependent increase in Pol II in gene bodies could be due to an increase in the rate of escape of Pol II from promoter locations to gene bodies. Consistent with this possibility, the FACT-dependent increase in pause release ratio fold change of highly paused genes is associated with an increase in RNA transcript abundance.

Finally, I hypothesized that the degree to which modified histones, or the chromatin architecture can be reorganized (disassembled and /or reassembled) by FACT would be dependent on the level of pausing of the gene. Our preliminary results suggest that FACT dependent changes in H3K36me3, H3K4me3 and H2Av are greatest on genes with low pausing compared to highly paused genes.

Chapter 5

Discussion

The histone chaperone, FACT was previously characterized in vitro as a factor that stimulates transcription elongation over chromatinized templates by perhaps displacing H2A-H2B (Belotserkovskaya et al., 2003; Orphanides et al., 1998). Taking a genome-wide approach we observed that not only is FACT necessary to maintain the chromatin landscape during transcription but also plays a significant role in regulating the interplay between promoter proximal Pol II pausing and transcription-coupled histone modifications in higher eukaryotes.

Promoter proximal pausing is now widely accepted as a regulatory step in the transcription of majority of active genes of eukaryotes (Jonkers and Lis, 2015). The role of FACT in regulating transcription initiation and elongation is not well understood in higher eukaryotes where transcription programs are complex. Based on the initial identification of FACT as an activity that enhances the ability of Pol II to elongate through nucleosomes on chromatin templates, I anticipated that FACT loss might make elongation, including the escape from promoter proximal pausing, less efficient. Data we have provided are indeed consistent with a role of FACT in regulating the efficiency with which Pol II can escape from promoter proximal pausing; however, and somewhat surprisingly, taken together the data presented in this thesis is most consistent with the idea that FACT depletion leads to an increase in the rate of Pol II release from promoter-proximal pausing. Using PRO-seq, we observed that the depletion of FACT leads to a reduction of transcriptionally engaged Pol II around promoters, accompanied by an increase in Pol II in gene bodies of many genes. Overall, the depletion of *Drosophila* FACT led to an increase in pause release ratio fold change, with the class of highly paused genes recording the

biggest increase. Since paused release ratio is a metric indicative of productive elongation (Chen et al., 2015a), we predicted an increase in differential gene expression for highly paused genes in the FACT depleted cells. Consistent with our prediction, we observed a significant increase in steady state RNA-seq among highly paused genes compared to moderately or lowly-paused genes (Figure 10D) following the depletion of SSRP1. These results suggest that the loss of FACT leads to an increase in productive elongation, occurring more frequently among highly paused genes.

Several factors may explain the possible involvement of FACT in maintaining promoter proximal pausing. Firstly, FACT-dependent changes in nucleosome occupancy or positioning may influence the efficiency of promoter proximal pausing. Our data suggests that depleting FACT subunits in S2 cells leads to small but reproducible loss of nucleosomes. It is conceivable that depletion of promoter proximal nucleosome(s) during repeated rounds of transcription of a gene could remove a nucleosomal block to pause release. If this is true, we might expect to be able to detect changes in nucleosome dynamics following FACT depletion using a more sensitive method known as “CATCH-IT” (covalent attachment of tags to capture histones and identify turnover) which involves the pulse labeling of newly synthesized histones and the subsequent purification of newly assembled nucleosomes (Teves et al., 2012).

Secondly, FACT-dependent changes in pause inducing (DSIF and NELF) or relieving (PTEFb) factors could modulate promoter proximal pausing efficiency. DSIF and NELF have been shown to play a crucial role in maintaining RNA polymerase II at paused regions (Wu et al., 2003; Yamaguchi et al., 2013). In order for RNA Pol II to escape the promoter proximal region, the kinase P-TEFb comprising of Cyclin T1 and CDK9 phosphorylates DSIF, NELF and the CTD Ser2 of the Pol II subunit Rpb1 to relieve the Pol II of the inhibition (Peterlin and Price, 2006). In addition, depletion of NELF and DSIF has been demonstrated

to alleviate pausing *in vivo*. If indeed FACT influences promoter proximal pausing by modulating pause inducing and relieving factors, it would be evident after performing ChIP-seq experiments of DSIF, NELF, PTEFb and Pol II ser2P in FACT depleted cells.

Finally, FACT has been reported to interact physically and functionally with transcription initiation and elongation modulators including Spt4/5 (Krogan et al., 2002; Squazzo et al., 2002) and the PAF complex- a protein complex implicated in promoter proximal pausing (Chen et al., 2015a; Yu et al., 2015). Notably, loss of PAF leads to increased pause-release coupled with increased CTD phosphorylation. Also, our data suggest that depleting FACT leads to increased pause release. Therefore, since FACT interacts with PAF, and depleting either PAF or FACT leads to increase in pause release, it is tempting to speculate that FACT might by an unknown mechanism modulate PAF function. If this is the case, one might predict that FACT depletion, like PAF depletion, might lead to changes in Ser2 phosphorylation and recruitment of Spt4/5, PAFc, P-TEFb or CDK12, ELL or other SEC components.

The transcription unit is marked by H3K4me3 at 5' ends (Barski et al., 2007; Schneider et al., 2004) and H3K36me3 within gene bodies (Wagner and Carpenter, 2012). Pol II CTD Ser5-P and Ser2-P contribute to the positioning of H3K4me3 and H3K36me3 respectively, since the Set1 H3K4 methyl transferase interacts with Ser5-P Pol II CTD (Ng et al., 2003) and the SET2 H3K36 methyl transferases interacts with Pol II CTD Ser2-P (Kizer et al., 2005; Sun et al., 2005). Several lines of evidence suggest that changes in the rate of elongation, and hence the length of time Pol II spends in a particular region of the transcription unit, can affect the distribution of transcription-coupled modifications. Similarly, patterns of H3K4 methylation are sensitive to elongation rate in yeast, being shifted in a 5' direction in cells expressing slowly elongating Pol II mutants, and in a 3' direction in cells expressing fast Pol II mutants (Soares et al., 2017). Based on these

observations, it has been proposed that the position of these marks can be altered by the length of time that H3K4 or H3K36 methyltransferases associated with Pol II phosphorylated on Ser5 or Ser2 of the Rpb1 CTD, respectively, spend at different regions of the transcription unit during multiple rounds of transcription (Fong et al., 2017; Soares et al., 2017).

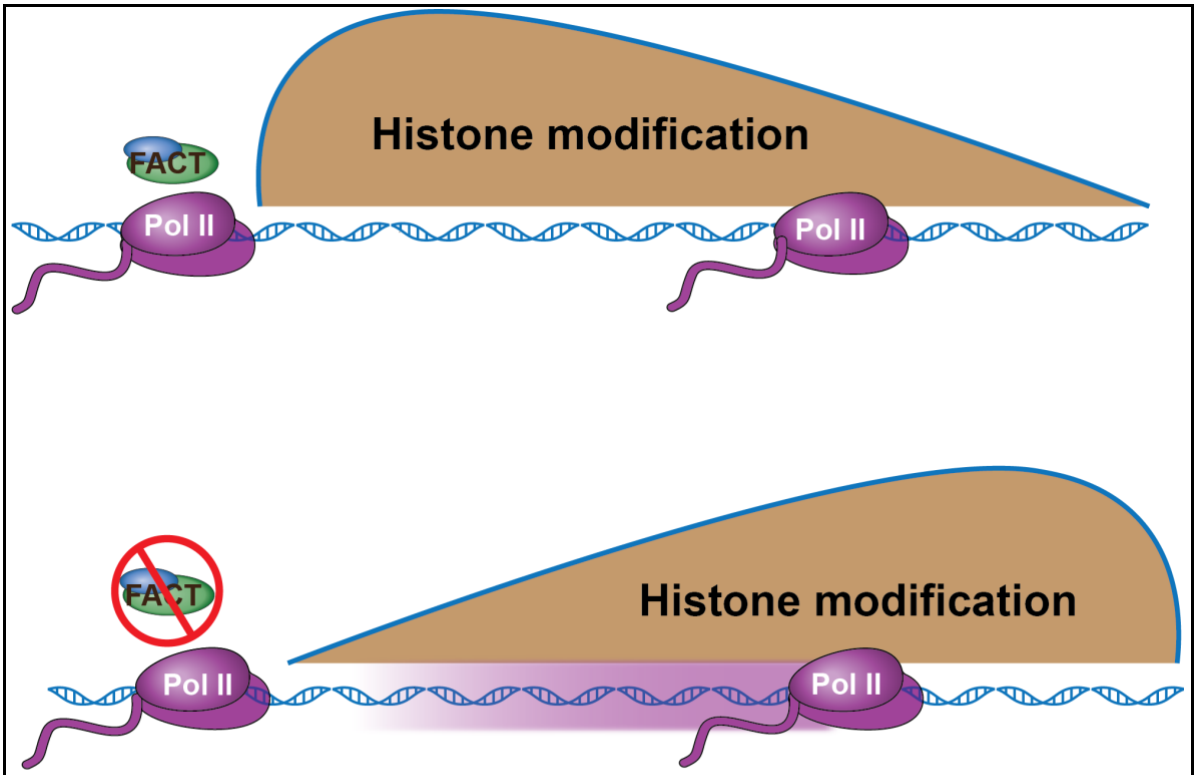

Figure 13. FACT contributes to the interplay between promoter proximal pausing regulation and transcription-coupled histone modifications.
In control cells (top panel), under normal Pol II transcription rates, histone modifications are co-transcriptionally deposited, where H3K4me3 occupies 5' ends and H3K36me3 occupies gene bodies. Upon depletion of FACT (bottom panel), Pol II moves quickly into productive elongation from paused sites through the relevant regions, elongation complex-associated methyltransferases would be allowed less time to place marks at any given place along the transcript, leading to the 5' to 3' shift in histone modifications (H3K4me3 and H3K36me3).

Therefore, our observation that H3K4me3 shifted 5' to 3' in the FACT deficient cells made us to speculate that the broadening of H3K4me3 may be associated with transcription elongation. In light of these observations it is tempting to speculate that decreased H3K4 methylation in the promoter-proximal regions of genes could be a consequence of reduced of promoter-proximally paused Pol II. Similarly, although we have not measured directly elongation rate throughout gene body in FACT-depleted cells, it is tempting to speculate that the relative loss of H3K4 trimethylation within promoters could be due to increased elongation rate due to FACT loss. To address this possibility in the future it would be of interest to perform measurements of Pol II elongation rates throughout the gene body. Such measurements can be made by treating cells with the initiation inhibitor triptolide or the P-TEFb inhibitor flavopiridol and following movement of Pol II along the template using techniques such as PRO-Seq, which measures the position of transcriptionally engaged Pol II.

H3K36me3, the co-transcriptionally modified histone, marks transcribed regions and correlates with transcriptional activity (Pokholok et al., 2005). Results from our experiments demonstrate that depletion of subunits of FACT leads to global changes in H3K36me3 in the transcribed region and a surprising increase beyond the poly-A+ site of many genes exemplified by 18w and Inos (Figure 7B). Since the changes in global H3 levels and micrococcal nuclease patterns were minimal in the FACT depleted cells, we could rule out the possibility that the massive changes in H3K36me3 is due to changes in bulk H3 patterns.

Our findings suggest that H3K36 tri-methylation may depend on subunits of the FACT complex since the depletion of FACT led to global changes in H3K36me3. However, since the human H3K36 methyltransferase SETD2 activity is reported to modulate FACT

recruitment in human cell (Carvalho et al., 2013), we propose that FACT and H3K36me3 and/or Set2 may have interdependent functions.

In addition, we noted that the co-transcriptionally modified histone H3K36me3 shifted 5’ to 3’ in the FACT depleted cells. It would be interesting to ask whether the FACT-dependent increase downstream of the Poly-A+ site is due to mislocalization and/ or changes in activities of set2 (H3K36me3 methyltransferase) and /or dKDM4A (a specific H3K36me3 demethylase). Toward this end, I have raised antibodies against this methyltransferase and demethylase, and these can be used in ChIP assays.

The shifts of histone modifications from 5’ to 3’ following the depletion of FACT is consistent with a direct role of FACT in proper localization of these marks. Alternatively, given that H3K4- and H3K36 methyltransferases interact with transcribing Pol II, it is not inconceivable that the 5’ to 3’ shifts in H3K4me3 and H3K36me3 could be as a result of changes in the distribution of Pol II across genes, perhaps due to differences in overall elongation rate or in the rate of escape of promoter proximally paused Pol II into more 3' regions of genes.

In summary, our results taken together favor the model (Figure 13) that FACT plays a critical role in maintaining promoter proximal pausing. Thus, depletion of FACT leads to increase in the rate of escape of Pol II from paused sites thereby shifting co-transcriptionally modified histones 5’ to 3’.

Chapter 6

Reference List

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat. Rev. Genet *13*, 720-731.

Allfrey, V.G., FAULKNER, R., and Mirsky, A.E. (1964). ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. Proc. Natl. Acad. Sci. U. S. A *51*, 786-794.

Antosz, W., Pfab, A., Ehrnsberger, H.F., Holzinger, P., Kollen, K., Mortensen, S.A., Bruckmann, A., Schubert, T., Langst, G., Griesenbeck, J.*, et al.* (2017). The Composition of the Arabidopsis RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. Plant Cell *29*, 854-870.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science *301*, 1090-1093.

Biswas, D., Yu, Y., Prall, M., Formosa, T., and Stillman, D.J. (2005). The yeast FACT complex has a role in transcriptional initiation. Mol. Cell Biol *25*, 5812-5822.

Black, J.C., Van, R.C., and Whetstine, J.R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol. Cell *48*, 491-507.

Blanchette, M., Green, R.E., Brenner, S.E., and Rio, D.C. (2005). Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila. Genes Dev. *19*, 1306-1314.

Blanchette, M., Green, R.E., MacArthur, S., Brooks, A.N., Brenner, S.E., Eisen, M.B., and Rio, D.C. (2009). Genome-wide analysis of alternative pre-mRNA splicing and RNA-binding specificities of the Drosophila hnRNP A/B family members. Mol. Cell *33*, 438-449.

Bonisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? Nucleic Acids Res. *40*, 10719-10741.

Brewster, N.K., Johnston, G.C., and Singer, R.A. (1998). Characterization of the CP complex, an abundant dimer of Cdc68 and Pob3 proteins that regulates yeast transcriptional activation and chromatin repression. J. Biol. Chem *273*, 21972-21979.

Brewster, N.K., Johnston, G.C., and Singer, R.A. (2001). A bipartite yeast SSRP1 analog comprised of Pob3 and Nhp6 proteins modulates transcription. Mol. Cell Biol *21*, 3491-3502.

Briggs, S.D., Bryk, M., Strahl, B.D., Cheung, W.L., Davie, J.K., Dent, S.Y., Winston, F., and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev. *15*, 3286-3295.

Bruce, K., Myers, F.A., Mantouvalou, E., Lefevre, P., Greaves, I., Bonifer, C., Tremethick, D.J., Thorne, A.W., and Crane-Robinson, C. (2005). The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. Nucleic Acids Res. *33*, 5633-5639.

Bruhn, S.L., Pil, P.M., Essigmann, J.M., Housman, D.E., and Lippard, S.J. (1992). Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin. Proc. Natl. Acad. Sci. U. S. A *89*, 2307-2311.

Burgess, R.J., and Zhang, Z. (2013). Histone chaperones in nucleosome assembly and human disease. Nat. Struct. Mol. Biol *20*, 14-22.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., and Workman, J.L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Carvalho, S., Raposo, A.C., Martins, F.B., Grosso, A.R., Sridhara, S.C., Rino, J., Carmo-Fonseca, M., and de Almeida, S.F. (2013). Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription. Nucleic Acids Res. *41*, 2881-2893.

Celic, I., Masumoto, H., Griffith, W.P., Meluh, P., Cotter, R.J., Boeke, J.D., and Verreault, A. (2006). The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. Curr. Biol *16*, 1280-1289.

Chen, F.X., Woodfin, A.R., Gardini, A., Rickels, R.A., Marshall, S.A., Smith, E.R., Shiekhattar, R., and Shilatifard, A. (2015a). PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. Cell *162*, 1003-1015.

Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., Rodriguez, B., Xi, Y., Xia, Z., Chen, X.*, et al.* (2015b). Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat. Genet *47*, 1149-1157.

Clark-Adams, C.D., Norris, D., Osley, M.A., Fassler, J.S., and Winston, F. (1988). Changes in histone gene dosage alter transcription in yeast. Genes Dev. *2*, 150-159.

Clemens, J.C., Worby, C.A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B.A., and Dixon, J.E. (2000). Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways. Proc. Natl. Acad. Sci. U. S. A *97*, 6499-6503.

Das, C., Lucia, M.S., Hansen, K.C., and Tyler, J.K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. Nature *459*, 113-117.

Di Bussolo, V., and Minutolo, F. (2011). Curaxins: a new family of non-genotoxic multitargeted anticancer agents. ChemMedChem *6*, 2133-2136.

Ding, Q., He, K., Luo, T., Deng, Y., Wang, H., Liu, H., Zhang, J., Chen, K., Xiao, J., Duan, X.*, et al.* (2016). SSRP1 Contributes to the Malignancy of Hepatocellular Carcinoma and Is Negatively Regulated by miR-497. Mol. Ther *24*, 903-914.

Finch, J.T., Lutter, L.C., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M., and Klug, A. (1977). Structure of nucleosome core particles of chromatin. Nature *269*, 29-36.

Fong, N., Saldi, T., Sheridan, R.M., Cortazar, M.A., and Bentley, D.L. (2017). RNA Pol II Dynamics Modulate Co-transcriptional Chromatin Modification, CTD Phosphorylation, and Transcriptional Direction. Mol. Cell *66*, 546-557 e543.

Formosa, T., Eriksson, P., Wittmeyer, J., Ginn, J., Yu, Y., and Stillman, D.J. (2001). Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN. EMBO J. *20*, 3506-3517.

Garcia, H., Fleyshman, D., Kolesnikova, K., Safina, A., Commane, M., Paszkiewicz, G., Omelian, A., Morrison, C., and Gurova, K. (2011). Expression of FACT in mammalian tissues suggests its role in maintaining of undifferentiated state of cells. Oncotarget *2*, 783-796.

Garcia, H., Miecznikowski, J.C., Safina, A., Commane, M., Ruusulehto, A., Kilpinen, S., Leach, R.W., Attwood, K., Li, Y., Degan, S.*, et al.* (2013). Facilitates chromatin transcription complex is an "accelerator" of tumor transformation and potential marker and target of aggressive cancers. Cell Rep *4*, 159-173.

Gariglio, P., Bellard, M., and Chambon, P. (1981). Clustering of RNA polymerase B molecules in the 5' moiety of the adult beta-globin gene of hen erythrocytes. Nucleic Acids Res. *9*, 2589-2598.

Gasparian, A.V., Burkhart, C.A., Purmal, A.A., Brodsky, L., Pal, M., Saranadasa, M., Bosykh, D.A., Commane, M., Guryanova, O.A., Pal, S.*, et al.* (2011). Curaxins: anticancer compounds that simultaneously suppress NF-kappaB and activate p53 by targeting FACT. Sci. Transl. Med *3*, 95ra74.

Gilchrist, D.A., Dos, S.G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540-551.

Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S.K., Collins, J.B., Li, L., Gilmour, D.S., and Adelman, K. (2008). NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes Dev. *22*, 1921-1933.

Guillemette, B., Bataille, A.R., Gevry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. (2005). Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. PLoS. Biol *3*, e384.

Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R.M., and Zhang, Z. (2007). Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science *315*, 653-655.

Hewish, D.R., and Burgoyne, L.A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem. Biophys. Res. Commun *52*, 504-510.

Hondele, M., Stuwe, T., Hassler, M., Halbach, F., Bowman, A., Zhang, E.T., Nijmeijer, B., Kotthoff, C., Rybin, V., Amlacher, S.*, et al.* (2013). Structural basis of histone H2A-H2B recognition by the essential chaperone FACT. Nature *499*, 111-114.

Hudson, M.E., Pozdnyakova, I., Haines, K., Mor, G., and Snyder, M. (2007). Identification of differentially expressed proteins in ovarian cancer using high-density protein microarrays. Proc. Natl. Acad. Sci. U. S. A *104*, 17494-17499.

Ikeda, Y., Kinoshita, Y., Susaki, D., Ikeda, Y., Iwano, M., Takayama, S., Higashiyama, T., Kakutani, T., and Kinoshita, T. (2011). HMG domain containing SSRP1 is required for DNA demethylation and genomic imprinting in Arabidopsis. Dev. Cell *21*, 589-596.

Iouzalen, N., Moreau, J., and Mechali, M. (1996). H2A.ZI, a new variant histone expressed during Xenopus early development exhibits several distinct features from the core histone H2A. Nucleic Acids Res. *24*, 3947-3952.

Jackson, J.D., Falciano, V.T., and Gorovsky, M.A. (1996). A likely histone H2A.F/Z variant in Saccharomyces cerevisiae. Trends Biochem. Sci *21*, 466-467.

Jackson, J.D., and Gorovsky, M.A. (2000). Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. Nucleic Acids Res. *28*, 3811-3816.

Jeronimo, C., Watanabe, S., Kaplan, C.D., Peterson, C.L., and Robert, F. (2015). The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. Mol. Cell *58*, 1113-1123.

John, S., Howe, L., Tafrov, S.T., Grant, P.A., Sternglanz, R., and Workman, J.L. (2000). The something about silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex. Genes Dev. *14*, 1196-1208.

Johns, E.W. (1967). The electrophoresis of histones in polyacrylamide gel and their quantitative determination. Biochem. J *104*, 78-82.

Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nat. Rev. Mol. Cell Biol *16*, 167-177.

Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096-1099.

Keller, D.M., and Lu, H. (2002). p53 serine 392 phosphorylation increases after UV through induction of the assembly of the CK2.hSPT16.SSRP1 complex. J. Biol. Chem *277*, 50206-50213.

Kemble, D.J., McCullough, L.L., Whitby, F.G., Formosa, T., and Hill, C.P. (2015). FACT Disrupts Nucleosome Structure by Binding H2A-H2B with Conserved Peptide Motifs. Mol. Cell *60*, 294-306.

Kemble, D.J., Whitby, F.G., Robinson, H., McCullough, L.L., Formosa, T., and Hill, C.P. (2013). Structure of the Spt16 middle domain reveals functional features of the histone chaperone FACT. J. Biol. Chem *288*, 10188-10194.

Kizer, K.O., Phatnani, H.P., Shibata, Y., Hall, H., Greenleaf, A.L., and Strahl, B.D. (2005). A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. Mol. Cell. Biol. *25*, 3305-3316.

Koman, I.E., Commane, M., Paszkiewicz, G., Hoonjan, B., Pal, S., Safina, A., Toshkov, I., Purmal, A.A., Wang, D., Liu, S.*, et al.* (2012). Targeting FACT complex suppresses mammary tumorigenesis in Her2/neu transgenic mice. Cancer Prev. Res. (Phila) *5*, 1025-1035.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865-868.

Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V.*, et al.* (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol. Cell *12*, 1565-1576.

Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J.F. (2002). RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol. Cell Biol *22*, 6979-6992.

Kusch, T., Florens, L., Macdonald, W.H., Swanson, S.K., Glaser, R.L., Yates, J.R., III, Abmayr, S.M., Washburn, M.P., and Workman, J.L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. Science *306*, 2084-2087.

Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950-953.

Kwak, H., and Lis, J.T. (2013). Control of transcriptional elongation. Annu. Rev. Genet *47*, 483-508.

Kwon, S.H., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. (2010). Heterochromatin protein 1 (HP1) connects the FACT histone chaperone complex to the phosphorylated CTD of RNA polymerase II. Genes Dev. *24*, 2133-2145.

Laskey, R.A., Honda, B.M., Mills, A.D., and Finch, J.T. (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. Nature *275*, 416-420.

Laskey, R.A., Mills, A.D., and Morris, N.R. (1977). Assembly of SV40 chromatin in a cell-free system from Xenopus eggs. Cell *10*, 237-243.

Latham, J.A., and Dent, S.Y. (2007). Cross-regulation of histone modifications. Nat. Struct. Mol. Biol *14*, 1017-1024.

Leach, T.J., Mazzeo, M., Chotkowski, H.L., Madigan, J.P., Wotring, M.G., and Glaser, R.L. (2000). Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster. J. Biol. Chem *275*, 23267-23272.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc *1*, 729-748.

Li, J., and Gilmour, D.S. (2013). Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. EMBO J. *32*, 1829-1841.

Li, J., Moazed, D., and Gygi, S.P. (2002). Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. J. Biol. Chem *277*, 49383-49388.

Li, Y., Zeng, S.X., Landais, I., and Lu, H. (2007). Human SSRP1 has Spt16-dependent and -independent roles in gene transcription. J. Biol. Chem *282*, 6936-6945.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Maas, N.L., Miller, K.M., DeFazio, L.G., and Toczyski, D.P. (2006). Cell cycle and checkpoint regulation of histone H3 K56 acetylation by Hst3 and Hst4. Mol. Cell *23*, 109-119.

Mahat, D.B., Kwak, H., Booth, G.T., Jonkers, I.H., Danko, C.G., Patel, R.K., Waters, C.T., Munson, K., Core, L.J., and Lis, J.T. (2016). Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nat. Protoc *11*, 1455-1476.

Malone, E.A., Clark, C.D., Chiang, A., and Winston, F. (1991). Mutations in SPT16/CDC68 suppress cis- and trans-acting mutations that affect promoter function in Saccharomyces cerevisiae. Mol. Cell Biol *11*, 5710-5717.

Mason, P.B., and Struhl, K. (2003). The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. Mol. Cell Biol *23*, 8323-8333.

Masse, J.E., Wong, B., Yen, Y.M., Allain, F.H., Johnson, R.C., and Feigon, J. (2002). The S. cerevisiae architectural HMGB protein NHP6A complexed with DNA: DNA and protein conformational changes upon binding. J. Mol. Biol *323*, 263-284.

Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature *436*, 294-298.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell *11*, 709-719.

Noma, K., and Grewal, S.I. (2002). Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast. Proceedings of the National Academy of Sciences of the United States of America *99 Suppl 4*, 16438-16445.

O'Donnell, A.F., Brewster, N.K., Kurniawan, J., Minard, L.V., Johnston, G.C., and Singer, R.A. (2004). Domain organization of the yeast histone chaperone FACT: the conserved N-terminal domain of FACT subunit Spt16 mediates recovery from replication stress. Nucleic Acids Res. *32*, 5894-5906.

Okuhara, K., Ohta, K., Seo, H., Shioda, M., Yamada, T., Tanaka, Y., Dohmae, N., Seyama, Y., Shibata, T., and Murofushi, H. (1999). A DNA unwinding factor involved in DNA replication in cell-free extracts of Xenopus eggs. Curr. Biol *9*, 341-350.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S., and Reinberg, D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. Cell *92*, 105-116.

Orphanides, G., Wu, W.H., Lane, W.S., Hampsey, M., and Reinberg, D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature *400*, 284-288.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell *4*, 281-300.

Ozdemir, A., Spicuglia, S., Lasonder, E., Vermeulen, M., Campsteijn, C., Stunnenberg, H.G., and Logie, C. (2005). Characterization of lysine 56 of histone H3 as an acetylation site in Saccharomyces cerevisiae. J. Biol. Chem *280*, 25949-25952.

Papamichos-Chronakis, M., Watanabe, S., Rando, O.J., and Peterson, C.L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. Cell *144*, 200-213.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol. Cell *23*, 297-305.

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E.*, et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Powell, W., and Reines, D. (1996). Mutations in the second largest subunit of RNA polymerase II cause 6-azauracil sensitivity in yeast and increased transcriptional arrest in vitro. J. Biol. Chem *271*, 6866-6873.

Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. Cell *123*, 233-248.

Ransom, M., Williams, S.K., Dechassa, M.L., Das, C., Linger, J., Adkins, M., Liu, C., Bartholomew, B., and Tyler, J.K. (2009). FACT and the proteasome promote promoter chromatin disassembly and transcriptional initiation. J. Biol. Chem *284*, 23461-23471.

Recht, J., Tsubota, T., Tanny, J.C., Diaz, R.L., Berger, J.M., Zhang, X., Garcia, B.A., Shabanowitz, J., Burlingame, A.L., Hunt, D.F.*, et al.* (2006). Histone chaperone Asf1 is

required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. Proc. Natl. Acad. Sci. U. S. A *103*, 6988-6993.

Richmond, T.J., and Davey, C.A. (2003). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D., and Klug, A. (1984). Structure of the nucleosome core particle at 7 A resolution. Nature *311*, 532-537.

Rougvie, A.E., and Lis, J.T. (1990). Postinitiation transcriptional control in Drosophila melanogaster. Mol. Cell Biol *10*, 6041-6045.

Rufiange, A., Jacques, P.E., Bhat, W., Robert, F., and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. Mol. Cell *27*, 393-405.

Safina, A., Garcia, H., Commane, M., Guryanova, O., Degan, S., Kolesnikova, K., and Gurova, K.V. (2013). Complex mutual regulation of facilitates chromatin transcription (FACT) subunits on both mRNA and protein levels in human cells. Cell Cycle *12*, 2423-2434.

Saunders, A., Werner, J., Andrulis, E.D., Nakayama, T., Hirose, S., Reinberg, D., and Lis, J.T. (2003). Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. Science *301*, 1094-1096.

Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat. Cell Biol. *6*, 73-77.

Sevilla, A., and Binda, O. (2014). Post-translational modifications of the histone variant H2AZ. Stem Cell Res *12*, 289-295.

Shi, Y., and Whetstine, J.R. (2007). Dynamic regulation of histone lysine methylation by demethylases. Mol. Cell *25*, 1-14.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu. Rev. Biochem *81*, 65-95.

Shimojima, T., Okada, M., Nakayama, T., Ueda, H., Okawa, K., Iwamatsu, A., Handa, H., and Hirose, S. (2003). Drosophila FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor. Genes Dev. *17*, 1605-1616.

Soares, L.M., He, P.C., Chun, Y., Suh, H., Kim, T., and Buratowski, S. (2017). Determinants of Histone H3K4 Methylation Patterns. Mol. Cell *68*, 773-785 e776.

Squazzo, S.L., Costa, P.J., Lindstrom, D.L., Kumer, K.E., Simic, R., Jennings, J.L., Link, A.J., Arndt, K.M., and Hartzog, G.A. (2002). The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. EMBO J. *21*, 1764-1774.

Stanlie, A., Aida, M., Muramatsu, M., Honjo, T., and Begum, N.A. (2010). Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for

DNA cleavage in class switch recombination. Proc. Natl. Acad. Sci. U. S. A *107*, 22190-22195.

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol. Cell Biol *22*, 1298-1306.

Stuwe, T., Hothorn, M., Lejeune, E., Rybin, V., Bortfeld, M., Scheffzek, K., and Ladurner, A.G. (2008). The FACT Spt16 "peptidase" domain is a histone H3-H4 binding module. Proc. Natl. Acad. Sci. U. S. A *105*, 8884-8889.

Sun, X.J., Wei, J., Wu, X.Y., Hu, M., Wang, L., Wang, H.H., Zhang, Q.H., Chen, S.J., Huang, Q.H., and Chen, Z. (2005). Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. J. Biol. Chem. *280*, 35261-35271.

Teves, S.S., Deal, R.B., and Henikoff, S. (2012). Measuring genome-wide nucleosome turnover using CATCH-IT. Methods Enzymol. *513*, 169-184.

Toney, J.H., Donahue, B.A., Kellett, P.J., Bruhn, S.L., Essigmann, J.M., and Lippard, S.J. (1989). Isolation of cDNAs encoding a human protein that binds selectively to DNA modified by the anticancer drug cis-diamminedichloroplatinum(II). Proc. Natl. Acad. Sci. U. S. A *86*, 8328-8332.

Tsunaka, Y., Toga, J., Yamaguchi, H., Tate, S., Hirose, S., and Morikawa, K. (2009). Phosphorylated intrinsically disordered region of FACT masks its nucleosomal DNA binding elements. J. Biol. Chem *284*, 24610-24621.

VanDemark, A.P., Blanksma, M., Ferris, E., Heroux, A., Hill, C.P., and Formosa, T. (2006). The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. Mol. Cell *22*, 363-374.

VanDemark, A.P., Xin, H., McCullough, L., Rawlins, R., Bentley, S., Heroux, A., Stillman, D.J., Hill, C.P., and Formosa, T. (2008). Structural and functional analysis of the Spt16p N-terminal domain reveals overlapping roles of yFACT subunits. J. Biol. Chem *283*, 5058-5068.

Wada, T., Orphanides, G., Hasegawa, J., Kim, D.K., Shima, D., Yamaguchi, Y., Fukuda, A., Hisatake, K., Oh, S., Reinberg, D., and Handa, H. (2000). FACT relieves DSIF/NELF-mediated inhibition of transcriptional elongation and reveals functional differences between P-TEFb and TFIIH. Mol. Cell *5*, 1067-1072.

Wade, J.T., and Struhl, K. (2008). The transition from transcriptional initiation to elongation. Curr. Opin. Genet. Dev *18*, 130-136.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nat. Rev. Mol. Cell Biol *13*, 115-126.

Weber, C.M., Ramachandran, S., and Henikoff, S. (2014). Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Mol. Cell *53*, 819-830.

Wittmeyer, J., and Formosa, T. (1997). The Saccharomyces cerevisiae DNA polymerase alpha catalytic subunit interacts with Cdc68/Spt16 and with Pob3, a protein similar to an HMG1-like protein. Mol. Cell Biol *17*, 4178-4190.

Woodcock, C.L., Safer, J.P., and Stanchfield, J.E. (1976). Structural repeating units in chromatin. I. Evidence for their general occurrence. Exp. Cell Res *97*, 101-110.

Wu, C.H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., and Gilmour, D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes Dev. *17*, 1402-1414.

Xiang, Y.Y., Wang, D.Y., Tanaka, M., Igarashi, H., Naito, Y., Ohtawara, Y., Shen, Q., and Sugimura, H. (1996). Expression of structure-specific recognition protein mRNA in fetal kidney and Fe-nitrilotriacetate-induced renal carcinoma in the rat. Cancer Lett. *106*, 271-278.

Xin, H., Takahata, S., Blanksma, M., McCullough, L., Stillman, D.J., and Formosa, T. (2009). yFACT induces global accessibility of nucleosomal DNA without H2A-H2B displacement. Mol. Cell *35*, 365-376.

Yamaguchi, Y., Shibata, H., and Handa, H. (2013). Transcription elongation factors DSIF and NELF: promoter-proximal pausing and beyond. Biochim. Biophys. Acta *1829*, 98-104.

Yu, M., Yang, W., Ni, T., Tang, Z., Nakadai, T., Zhu, J., and Roeder, R.G. (2015). RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. Science *350*, 1383-1386.

Yuan, J., Pu, M., Zhang, Z., and Lou, Z. (2009). Histone H3-K56 acetylation is important for genomic stability in mammals. Cell Cycle *8*, 1747-1753.

Zhang, H., Roberts, D.N., and Cairns, B.R. (2005). Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. Cell *123*, 219-231.

Zhang, W., Zeng, F., Liu, Y., Shao, C., Li, S., Lv, H., Shi, Y., Niu, L., Teng, M., and Li, X. (2015). Crystal Structure of Human SSRP1 Middle Domain Reveals a Role in DNA Binding. Sci. Rep *5*, 18688.